DOCUMENT RESUME

ED 335 690 CS 212 985

AUTHOR Lofqvist, Gert
TITLE The IEA Study of Written Composition in Sweden.
INSTITUTION International Association for the Evaluation of Educational Achievement, Stockholm (Sweden).; Lund Univ. (Sweden). Malmo School of Education.
SPONS AGENCY National Swedish Board of Education, Stockholm.
REPORT NO ISBN-91-7900-901-8
PUB DATE 90
NOTE 168p.; For an earlier report, see ED 317 997.
PUB TYPE Reports - Research/Technical (143) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC07 Plus Postage.
DESCRIPTORS Elementary Secondary Education; Foreign Countries; *Writing (Composition); Writing Instruction; Writing Research; *Writing Skills
IDENTIFIERS *IEA Written Composition Study; International Assn Evaluation Educ Achievement; *Sweden

ABSTRACT

This book describes and discusses the International Association for the Evaluation of Educational Achievement (IEA) study of writing in Sweden. The chapters are as follows: (1) Introduction; (2) Content and Planning of the Study; (3) Schools, Teachers and Pupils in Sweden 1985; (4) The Teaching of Writing; (5) Compositions and Scoring; (6) Relations between Background Factors and the Students' Writing Capacity; (7) Content of the Compositions; and (8) Pedagogical Consequences. Sixty-six references are attached. (SR)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED335690

Gert Löfqvist

The IEA Study of Written Composition in Sweden

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

a. Bjerstedt

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it

☐ Minor changes have been made to improve reproduction quality

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

BEST COPY AVAILABLE

C5

The International Association for the Evaluation of Educational Achievement
IEA Written Composition

Gert Löfqvist

The IEA Study of Written Composition in Sweden

CS212985

Lund 1990

This dissertation is no.93 in the series
STUDIA PSYCHOLOGICA ET PAEDAGOGICA – SERIES ALTERA

© Gert Löfqvist
ISBN 91-7900-901-8
Printed in Sweden
Studentlitteratur
Lund 1990

Contents

Acknowledgements

1 Introduction 7
- 1.1 The international organization IEA 7
- 1.2 The study of writing 7
- 1.3 The Swedish study 8

2 Content and planning of the study 11
- 2.1 Aims of the project 11
- 2.2 Domain of writing - a theoretic and empiric background 12
- 2.3 Tasks and questionnaires 21
  - 2.3.1 Motivations for the choice of writing tasks 21
  - 2.3.2 Writing tasks 27
  - 2.3.3 The questionnaires 31
- 2.4 Target groups, planning and realization 32

3 Schools, teachers and pupils in Sweden 1985 35
- 3.1 Schools 35
  - 3.1.1 Material standard 35
  - 3.1.2 Cultural standard 36
  - 3.1.3 Internal work 37
- 3.2 Teachers 38
  - 3.2.1 Age and education 39
  - 3.2.2 Working conditions 39
- 3.3 Students 40
  - 3.3.1 Academic and vocational ambitions 40
  - 3.3.2 Immigrant students 40
  - 3.3.3 Supply of books and papers 41
  - 3.3.4 Interest in reading 41
  - 3.3.5 Homework 42
  - 3.3.6 TV and radio 43
  - 3.3.7 Family climate 44
  - 3.3.8 Students' world of ideas 44

4 The teaching of writing 46

4.1 Aims and principal items 46

4.2 Writing practice and preparations for writing 47

4.3 Writing tasks 48

4.4 Realization and scoring 50

5 Compositions and scoring 52

5.1 Method of scoring - a theoretic and empiric background 52

5.2 The scorers' guide 57

5.3 Evaluation results 64

5.3.1 Model of scoring 66

5.3.2 Training and control of scorers 68

5.3.3 Results of the main study 81

5.3.4 Language analysis and written achievement 111

6 Relations between background factors and the students' writing capacity 121

6.1 Factors that influence the students' writing capacity - a theoretical model 121

6.2 Structuring of information 123

6.3 Partial Least Squares (PLS) - a short presentation 133

6.4 Relations within background factors and the written achievements within different age groups 133

6.5 Relations within the same group for different written tasks 142

7 Contents of the compositions 147

8 Pedagogical consequences 150

8.1 Teaching of writing 150

8.2 Reliability of scoring 156

8.3 Scoring method 159

References 161

Acknowledgements

The assistance of a lot of persons made this dissertation possible. I owe them many thanks.

Foremost I have got impulses and ideas from them who conducted this international study of written composition and who worked as international coordinators. I want naturally to thank the members of the Swedish National Committee and especially its chairman Inger Marklund, head of the planning unit at the Swedish Board of Education, for their help.

Also I want to offer my sincere thanks to all the headmasters, teachers and students who have readily participated in the project. Without their help this study could never been carried out.

Many teachers and colleagues at Malmö School of Education helped me in collecting the material, in classification, in scoring the compositions, and in computeranalyses. I owe a debt of gratitude to them all.

I am very grateful especially to some persons for their assistance. Professor Ebbe Lindell interested me in this study, and he has patiently spent much time to discuss the problems of planning and handling the information. Professor Horst Löfgren has helped me with the analyses and has given me statistical valuable advice. Lecturer Sven Fjelner has with his deep knowledge in English assisted in translations to English on several occasions, also with this dissertation. With the greatest accuracy Siv Karlén has, among other things, typed most of my reports including this one.

Finally I thank the Swedish Board of Education for the financial support to this study.

Malmö, November 1989.

Gert Löfqvist

1 Introduction

1.1 The international organization IEA

At the end of the 1950s a number of scholars from a dozen countries came together to discuss empirically oriented comparative research problems. The organization IEA was created in 1961. The purpose was to bring about international co-operation in order to be able to make comparisons between different national educational systems.

A minor preliminary study with twelve countries taking part was carried out 1959-1962. It was followed by the first mathematics investigation in 1964 and two years later the so-called six-subject study started. It included science, reading comprehension, literature, French and English as foreign languages and civics.

A second investigation in mathematics and science has been completed during the 1980s. In the course of this decade the work on the investigation of writing has been going on.

Roughly 10-20 countries have participated in the various investigations but only Sweden has taken part in all of them.

1.2 The study of writing

The IEA Study in Written Composition started in 1980 after a few years of preliminary discussions. In 1980 an inquiry was made to professors Gunnar Hansson and Ebbe Lindell about Swedish co-operation in the project. The National Board of Education in principle assumed a positive attitude towards participation but at that time found it impossible to finance Swedish co-operation with regard to other national and international commitments.

The work within IEA concentrated on building a model for writing that might be the basis of the construction of tasks for essays as well as the compilation of an evaluation system that would make international comparisons of the student achievements possible. The model for writing is described in chapter 2 and the evaluation system in chapter 5.

An extensive collection of questionnaires concerning the cultural and material level of the participating countries, their educational systems and school organization had been worked out through the international secretariat. Also questionnaires were constructed to be used by the students, their teachers of the mother tongue and the headmasters of the schools that took part in the investigation. An extensive background material would thus be collected to make it possible to explain variations in student achievements. Much preliminary work was carried out during the first years consisting of the trying out of tasks and questionnaires as well as organizing manuals etc. As Sweden did not take part from the beginning, we have not joined in this preparatory work.

The participating countries are, in addition to Sweden, Chile, England, Finland, Indonesia, Italy, the Netherlands, Nigeria, New Zealand, Thailand, West Germany (Hamburg), Hungary, the USA and Wales. A few other countries have been connected with the project during certain parts of the investigation.

1.3 The Swedish study

In the autumn of 1983 scholars and representatives of the National Board of Education discussed Swedish participation. The debates ended in an agreement that the compositions within the project should be examined and tested out and possibly modified according to Swedish conditions. It was decided that I should be the Swedish representative at the project conference at Urbana in the USA during the spring of 1984. It was agreed that the final decision on Sweden's taking part in the project should wait until this conference had finished.

Each country that takes part in an IEA study elects a national committee to serve as an advisory body. In the Swedish committee for the study in Written Composition were included from 1984 Inger Marklund, head of the planning unit at the Swedish Board of Education, Karin Dahl, head of a department at the Swedish Board of Education, professor Gunnar Hansson, senior lecturer Kent Larsson, professor Ebbe Lindell, school inspector Gert Löfqvist, professor Ingemar Wedman, senior lecturer Margareta Westman, and Britt Wilson Lohse, head of a department at the Swedish Board of Education. Professor Lindell has been scientific leader and has been responsible for international contacts, and I have been responsible for arrangement and reporting.

My work has consisted of some routine duties and also such independent contributions as those that are accounted for in this dissertation, some of which have been published earlier in other reports. Among the latter duties the following should be mentioned, and they make up the themes of the dissertation.

To describe the conditions in Sweden for the investigation of writing

During the autumn of 1983 a pilot study was planned (Löfqvist, 1988a). It aimed at testing out the written compositions and collecting material for conferences in the USA and Italy on evaluation problems. Certain changes were made in the wording of the written tasks. When the study had been completed the national committee discussed Swedish participation. Reasons for and against such participation were given (Löfqvist, 1988a, p 31). A decision in favour of participation was arrived at in May 1984.

To adapt the questionnaires to Swedish conditions

The questionnaires to students, teachers and headmasters have been adapted to make them suited to Swedish school conditions. Certain questions have been left out, others have been changed from the point of view of content and some have been added (Löfqvist, 1988a, pp 57-67). Schedules for the categorization of questions with free answers have been prepared (pp 68-79).

To give a description of writing education

The questionnaire answers have been revised and structured. In chapters 3 and 4 there are reports on school conditions and the teaching of writing in Sweden. A comparison between Swedish and foreign teachers' background data and ecucational practice has been described separately (Löfqvist, 1988c).

To score the written compositions

The scoring model has been worked out within the scope of the national freedom. The guide for the scorers has been used in broad outline but adapted to Swedish conditions. Some of the members of the national committee have taken part in the national adaptation. The remaining parts of the account in chapter 5 refer exclusively to Swedish conditions.

To seek causes for variations in student achievements

Preliminary discussions within the international project have been conducted between the participating countries on methods for the identification of factors that can explain variations in student achievements. The report in chapter 6 of the structuring of background information and of the causal analyses has however been a national concern.

2 Content and planning of the study

2.1 Aims of the project

The teaching of writing or what is usually included in the word composition is a field of research that requires both theoretic and empiric contributions. Writing has not been well defined internationally. Opinions differ as to what should be considered to be good or satisfactory solutions of written assignments. Different countries vary as to the importance they attach to different aspects of writing, the planning of a writing methodology and evaluation systems. It is therefore natural that there should not exist any writing tasks to be applied internationally. Nor can different countries be expected to agree on what to consider as criteria of good achievements in writing or what evaluation methods to be used.

The first problem to solve within the project was therefore to analyse and define the concept of writing and especially writing at school. This was considered necessary before the tasks were constructed. In this connection it has also been important to work out an evaluation system that could be used internationally. First work was carried out within the project aiming at getting the evaluation so reliable that comparisons between students achievements in different countries would be possible.

The second task is to make descriptions of schools, teachers and pupils concerning conditions that have no direct connections with the teaching of Swedish but may still be of importance to the results in writing and elucidate and describe the contents of the teaching of writing, the methodology and evaluation methods for each country and after that make comparisons between different countries.

A third purpose is for each country and group of students to describe the student achievements of the different tasks. In this connection the aim is to identify such factors for each country both within and among groups of students as can be related to student, teacher, school and home and can explain variations in performances.

The term written composition almost corresponds to the word essay writing. In Sweden we usually used the words "essay" and "essay writing" up to 1972 in the compulsory school, whereas they have lived on in the upper secondary school. In the new compulsory school curriculum, however, one wished to get away from the test situation of the old school composition by replacering the term "essay" with the vague concept sketch or outline. This was completed with expressions such as free writing, free written narrative, written report, writing, etc.

Even if the intention was probably to mark a change from long written tests to shorter and simpler written exercises in Swedish schools, the idea was still there that the students, at least in the upper level of the compulsory school, should have an opportunity to write during a long, continuous period of time and that their writings should be marked according to a grading system given in the curriculum. In the 1980 curriculum for the compulsory school a different methodological aim and direction of the teaching of writing has been recommended, which describes writing as a long process (Löfqvist, 1989a). The terms essay or composition have not been reintroduced, however.

For practical reasons the words essay and composition are used in most cases in this thesis. In references to other investigations it seems natural to do so, if the writers themselves use these terms. But the shorter written tasks in the study can hardly be called compositions in the accepted sense of the word. However, it has turned out to be too complicated in the discussions to try and separate the tasks that could be marked as compositions from the shorter written tasks. Task essay, essay and composition are therefore here used as almost synonymous expressions.

2.2 Domain of writing - a theoretic and empiric background

In an international study it is necessary for the participating countries to use a common terminology and to agree on the processes of thought that constitute among other things the basis of the measuring instruments. As has been earlier pointed out it was therefore necessary first of all to make up an analysis of concepts.

Takala (1983), the international secretary of the project, in his survey of the present research on writing has tried to distinguish different trends. One of them is focused on the importance of the abilities to read and write on the individual's cognitive and social development, another on the relations between speech and writing, still another on the written product with studies of textual structures and their grammatical, narrative and argumentative construction. Other fields of research have a pedagogic trend towards teaching methods, curricula and matters of evaluation.

According to Takala writing is usually considered as a communicative process between the writer and the reader. He refers among others to Collins and Gentner (1980), who identify four general and essential principles of all written communication. Comprehensibility is important and can among other things be facilitated by means of exemplifications. The demand for comprehensibility is especially called for in analysing texts. But even if the writer has managed to make the text easy to understand, he must also make it interesting. Collins and Gentner mention different methods to motivate the reader not to stop reading. The purpose of a text is often not only to explain or relate something but also to convince the reader of how true, important, authentic etc. the written storry is, and therefore the text ought to be argumentatively oriented. The most essential demand, however, seems to be to work out the text in such a way that it is easy to memorize. These general demands on writing are related to structure, style, genre and contents of the texts.

Vähäpassi (Takala & Vähäpassi, 1983; Vähäpassi, 1988) also finds writing action oriented in the social perspective, a communication directed from the writer towards the reader. The action aims at transferring in one form or another contents to the reader. According to Grice (1975) the writer as well as the reader have responsibilities that for the writer are partly parallel to Collins' and Gentner's demands on written communication. The writer must try to write a clear, relevant, truthful, informative, interesting and easily memorized text, whereas it is the reader's duty to read and interpret it in consideration of the writer's supposed intention.

In modern Swedish research above all Kent Larsson (1984) has discussed language and writing abilities in his dissertation. He means that in the study of language proficiency and also writing ability the starting-point should be taken in the concept of action rather than in the narrower concept of structure.

> "Language will to a rather large extent appear as a procedural or operative knowledge by means of which the individual creates order and significance in what goes on around him and as thus the central factor in the human activity which implies acquiring, organizing and using knowledge in a social community based on comunication. In this perspective the linguistic development is to be looked upon as a constructional process in which the individual himself by working together with other people and by handling the surrounding reality builds up a capacity for acting, an ability to take action that constitutes his language proficiency." (p 22)

In her discussion of the theories of the function of language of different scholars Vähäpassi especially takes Jacobson's (1960) theory as her startingpoint. Jacobson elaborated Bühler's model on the relations between the linguistic expression the sender, the receiver and the surrounding circumstances (Bühler, 1934). According to Jacobson each of the six factors in figure 1 determines different language functions.

CONTEXT
MESSAGE
ADDRESSER .... ADDRESSEE
CONTACT
CODE

Figure 1. Jacoboson's view of the functions of language (from Vähäpassi, 1988)

The emotive function is an expression of the writer's feelings and experiences. The sender thus determines the function. The message can be wordy and unstructured, perhaps as the receiver knows the writer well and the message is interpreted correctly because of their common experiences.

The function that principally aims at keeping contact between the sender and the receiver is by Jacobson called phatic. The content of the linguistic expression is unimportant but must be uncontroversial, for example opinions on the weather, as the intention is to keep the social contact open. Christmas and New Year greetings and postcards are characteristic types of text. In the conative function the receiver is in the limelight. The purpose is to change his behaviour, thinking and actions, and the message is adapted to the receiver's knowledge and attitudes. If the intention is a linguistic development, the lexical and syntactic code, the function is called metalingual. When the message determines the direction, Jacobson talks about the poetic function of language. It is then not only limited to poetry, and the writer does not only concentrate on the message but also on the code. Last of all the context plays a dominating part in the referential function when knowledge is to be acquired or presented, when opinions and ideas are to be analysed or when events are to be described or related.

We also meet with the view that all linguistic activities are a form of international action (Westman, 1980). With reference among others to Allwood (1976) she emphasizes that all that is written has a special purpose. It is really the intention that distinguishes any course of events from an action. What has been written is also always meant for a reader or some readers.

But Vähäpassi does not look upon writing at school only from a social angle aiming at purpose and receiver, but also from a cognitive point of view. Takala (1983) also discusses models of writing as a cognitive process and refers to Hayes and Flower (1980), Collins and Gentner (1980), Bereiter (1980) and others. The first-mentioned have presented a model intended to describe the intellectual process when professional writers make reports. The writing process contains three important phases, namely planning, translating into a suitable linguistic form and reviewing. Each of these has routines on lower levels.

According to Vähäpassi the intellectual levels in the writing process constitute a complementary dimension to the purpose. She means that the intention of writing is accentuated in many models of written work, but that there is a tendency to neglect the intellectual process. The connection between thinking and writing is emphasized but the relation between them is not developed in a way that would be applicable in the intellectual demands on written tasks:

1. To be able to receive, recognize and remember information:
   the writer reproduces events or facts, for example copies, quotes, notes.
2. To be able to analyse and structure reality:
   the writer for example relates, describes, explains, sums up.
3. To evaluate and recreate information:
   the writer widens, invents or creates a reality, for example analyses, explains, argues, creates a new possible world.

Vähäpassi concretized her ideas in a classification of written tasks at school. While constructing it she aimed at including such factors as the demands on the purpose of what is being written and the direction towards the readers as well as what is expected of the cognitive efforts related to topic and contents, partly also the linguistic and rhetorical demands on the type of text.

In her discussion of types of text Vähäpassi refers to Brewer (1980), who distinguishes between three ways of writing, namely describing, narration and explaining. These types of text differ from each other in the cognitive structure. The descriptive type tries to transpose a field of perception into a verbal form with a chiefly visual-spatial cognitive structure. In the relating type of text the events of a series are related to each other by means of causal and thematic couplings. The explanatory text comprises inductive, deductive, classifying and comparing structures. Vähäpassi believes that the development of an individual's writing ability is closely connected with the intellectual development. Therefore she finds it reasonable not to differentiate between the way of writing and the processes of thought. The following grouping of types of texts, chiefly based on a rhetorical tradition, may therefore be usable.

1. Documentative discourse:
   The material gets little or no revision.

2. Reportorial discourse:
   a) narrative: events and experiences
   b) descriptive: static conditions or dynamic processes,
   c) explanatory: logical reports based on comparisons, groupings or causal arguments.

3. Exploratory discourse:
   a) interpretive (expository/argumentative/persuasive):
   summaries, reports, articles etc. based on revised material (induction, deduction, comparisons etc.)
   b) literary discourse.

Vähäpassi's classification of writing at school is represented in figure 2.

Dominant intention/ Purpose	Cognitive Processing	Reproduce		Organize/Reorganize		Invent /Generate	
	Primary Content Primary Audience	Facts	Ideas	Events	Visual images facts,mental states,ideas	Ideas ,mental states alternative worlds	
To learn (metalingual)	Self	copying, taking dictation		retell a story (heard or read)	note, resume, summary, outline, paraphrasing	comments on book margins metaphors analogies	
To convey emotions, feelings (emotive)	Self, others	stream of consciousness		personal story personal diary, personal letter	portrayal	reflective writing personal essays	The traditional literary genres and modes can be placed under one or more of these five purposes
To inform (referential)	Others	quote	fill in a form	narrative report news, instruction, telegram, announcement, circular	directions, description, technical description, biography, science report/ experiment	expository writing definition academic essay/article book review commentary	
To convince/ persuade (conative)	Others	citation from authority/expert		letter of application statement of personal views opinions	advertisement, letter of advice	argumentative / persuasive writing editorial critical essay/ article	
To entertain, delight,please (poetic)	Others	quote poetry and prose		given an ending create a story create an ending retell a story	word portrait or sketch	entertainment writing parody,rhymes	
To keep in touch (phatic)	Others	greeting card		postcard	personal letter	humorous greeting	

Figure 2. Classification of writing tasks at school (from Vähäpassi, 1988)

In the classification one dimension refers to the main purpose of the students' writing. The other consists of the theoretical and intellectual demands that written work confronts the students with. With these main dimensions she connects the basic reader and the basic contents. Below the intellectual levels can be found in the different boxes the documentary, reporting and examining types of texts. The classification indicates the difference between types of text and purposes of writing, two notions that in the rhetorical tradition have in most cases been considered synonymous.

The dimensions Dominant intention/Purpose is being described, according to Jacobson's functions, in the left-hand column in six ways: to learn, to convey emotions, feelings etc. The second dimension, Cognitive Processing, is grouped into the categories Reproduce, Organize/Reorganize and Invent/ Generate with an increase of the demands on an intellectual effort from left to right.

In the boxes in the figure it is possible to put in different types of writing determined by the main dimensions. Such a classification has been made of a number of more or less common types of writing, but this exemplification is not meant to be complete. Nor is it final in the sense that a certain type of writing can only be referred to one definite box. A summary may for exemple be written with the purpose of teaching, but it could also aim at informing or convincing. The table rather shows how diversified the field of writing at school is or can be.

A few Swedish studies have treated writing at school. The most extensive of these are "Free writing at school" (FRIS), which deals with the intermediate stage (Lindell, 1980), and the project "Writing syntax", a study of the written language of upper secondary school students (Hultman & Westman, 1977). They both differ from the IEA study in such a way that they are essentially investigations of the linguistic usage of the students. In the FRIS project there are investigations of the language of the students and its development during the intermediate stage according to linguistic measures and a validation of the linguistic phenomena according to essay marks (Martinsson, 1978), further studies of differences in linguistic ability owing to sex and social group (Lundquist, 1978) and of creativity in writing, related among other things to the linguistic measures, finally experiments with varied types of stimulus (Pettersson, 1978a, 1978b). The project has observed 191 students, living in different socio-economic areas, from grade 4 to grade 6. In different contexts from now on there are reasons for me to make comparisons between the FRIS project and the IEA study.

Hultman and Westman have in their study "Senior high school Swedish" drawn material from the students' compositions in the central tests of 1970. The language of 151 compositions, all dealing with the written task "Family and marriage once more", is chiefly compared with every day Swedish, represented by the genres broschures, newspaper articles, schoolbooks and debate texts. The comparisons between the language of upper secondary school students and the language of adults, between student compositions with different marks and between the sexes are carried out with regard to word frequences, the use of parts of speech, syntactic and graphic differences. The writers elucidates positive ingredients in the students' language but also emphasize deficiencies and give continuous advice as to how the students might improve their language. It will also have occasion to come back to this investigation.

Among Swedish scholars within the field of writing Kent Larsson (1984) has been mentioned earlier. His research aims at getting through to a deeper knowledge of the writing ability and its development among students. One of the principal questions has reference to the relation between writing ability and social conditions. Larsson finds a strong connection between written achievement expressed in marks and socio-economic status. A second principal question is the relation between written achievement and productivity. I will return to Larsson's results in what follows.

An important aspect has furthermore been to describe what the writer is doing when writing. Larsson means that through a series of operations in five levels the writer builds up a textual world. The different levels form the basis of his empiric investigation of students' written achievements. The writing ability is related to spelling (level 1), word knowledge (level 2), textual operations on a micro level (level 4), measured with two tests consisting of sentences and paragraphs to arrange in the correct order, as well as textual operations on a macro level (level 5) which mean dividing into paragraphs understructured texts. Level 3 bears reference to the demands on sentence-construction, but it is not tested. The students' achievements on different levels are related to reading performances, grade and socio-economic status. A number of written language proficiency profiles are introduced to illustrate the relation between word level and the textual levels. The results are discussed in comparison with the IEA investigation in the following. Larsson used a sample of totally 351 students from grades 7 and 9 of the compulsory school and grades 2^2 and 2^3 of the upper secondary school.

In another investigation at Uppsala university with a selection of the compositions from Larsson's study Garme (1988) has analysed how students work when writing. In a free choice of themes in the compositions school descriptions were predominant. Older writers were more interested in motives dealing with human relations. Garme describes how the students set to work at their assignment after a very short preparation either by tackling the topic right away by giving the story a frame consisting of a linking introduction and conclusion, by an indirect attachment to the task, or by not using the stimulus material at all. The most common model is to use a framework, a method which was also favoured according to a questionnaire used by Garme with another group of students.

The students can use different ways of defining or marking off the written task by headlining, summarizing or trying to adapt themselves to their readers. Direct speech seems to be found above all among the best compositions. Garme has also examined the attitude of the writer towards the text and estimations in the texts. In different connections she makes comments that ought to be paid attention to in the teaching of writing.

A training program for writing in the upper secondary school (the STIG project) has aimed at comparing the development of writing when two different methods of writing practice were applied (Pettersson, 1982). In one group students in grade 1 of the upper secondary school were trained by means of regular teaching of writing. In a conventional way they had an opportunity to practise relating, describing, summarizing, explaining and arguing. The other group practised transformations and concentrations of the language on a sentence level. Functional writing practice was thus contrasted with formal training. Pettersson found after half a year's training of the students that the formal method had led up to shorter compositions than the functional method but that both of them gave the same variation of words. The language was more like adult language after the formal training.

Grundin (1975) has mapped out the development of the writing ability through the school years above all using reading tests but also with a test in productive writing. The estimate of the written achievements consisted of counts of the number of words and of long words in the compositions, methods of calculation that have turned out to be very reliable both in earlier investigations (Lindell, 1980; Hultman & Westman, 1977) and in this IEA study. Grundin found a continuous increase through the whole of the compulsory school of the average number of long words. With increasing age there was thus an enhancement of vocabulary in the students' compositions. No corresponding continuous development of the number of words could be verified. The numbers of words increased rapidly at the junior level but became weaker at the intermediate level and then increased again continuously. Grundin defines a minimum level for what he calls functional reading and writing proficiency, which each student ought to reach before the end of the nine-year compulsory school. The minimum level for writing is the skill that the average student in grade 6 has achieved, and about 15% of the students in grade 9 do not get as far as that.

Out of the Swedish research contributions in the field of writing quite a few have points in common with the IEA study. In the following I will return to these for discussions and comparisons with this study but also to other investigations, which have not been mentioned here, for example Eneskär's (1988) report about the results on the Swedish standardized achievement tests.

2.3 Tasks and questionnaires

2.3.1 Motivations for the choice of writing tasks

The classification of writing tasks which was presented in the previous section has been the basis of the construction of the writing tasks of the project. It has however not been possible to create different tasks entirely with the help of the model. It has also been necessary to pay attention to the curricula of the participating countries writing traditions. It became apparent that the teaching of different writing tasks varied between the countries. Even within countries writing is practised in different ways due to the construction of the school system. In countries with centrally designed curricula and a homogeneous school system the situation is different from the one in countries with a strongly decentralized school.

At the same time as there was an ambition to consider different writing traditions, there was also an aim at giving a variety of tasks of different sorts. Partly they should be designed in such a way that the students had to write with a special intention, partly they should be on varying abstraction levels and thus be intellectually differently demanding. By filling as many boxes as possible in the model validity was created for the test instrument.

The choice of writing tasks has been made guided by what was already known about what students of different age groups are able to master. There was an ambition to make such a selection that it would not favour or be unfair to students from countries with different school systems. Naturally it has been impossible to avoid that students should be given tasks that they had not earlier had any practice of, as countries of different character and on varying cultural and material levels take part in the study.

The construction of written tasks, for example in the form of examination papers, varies a lot in different countries. As it was important to find a solution balancing between too general, vague expressions on one hand and limited detail instructions with stereotyped compositions as a result on the other, a pilot study was made in the schools of the participating countries. This study showed among other things if certain tasks were decidedly unsuitable to be used and what kinds of stimuli, for example texts and pictures, that functioned best.

In the wording of the tasks instructions were put in partly about the principal purpose, for example to inform the imaginary reader, partly about the type of writing, for example summary or letter, and also about the type of discource for example telling a story or giving a description, and furthermore about content. In addition to this, evaluation criteria and writing time were stated.

Baker (1982) found that instructions to the students, for example telling them to describe, argue or analyse or to exemplify, showing them that they are expected to write a certain number of pages, are essential to the quality of the compositions. Also the way a stimulus is given for the writing of a composition is of great importance. According to Bereiter and Scardamalia (1981) the use of varying stimuli, for example pictures, texts, music etc., tends to result in compositions that are richer both as to contents and language, than if only a headline has been given. Efforts have been made during the construction of the tasks to use such experiences but a certain amount of moderation has also been necessary concerning the variation of stimuli and information about the assignment. In many of the participating countries the students get very few instructions about the tasks, and there was a wish not to deviate too much from the usual writing arrangements.

As it was impossible to demand so much time from teaching that all students wrote all the different tasks, it was decided that the tasks should be rotated, so that each student did only certain tasks. How this rotating was carried out and how the tasks were distributed to the different groups of students will be descibed in section 2.4.

These considerations led up to a construction of tasks with a distribution to the boxes in the model for writing according to figure 3. Some of the tasks were compulsory, and they have been marked with an asterisk. Sweden made use of these tasks only. They are described in the following. The letters, A, B and C refer to different student groups. They are defined in 2.4.

	Primary Cognitive Demand		
Dominant intention/ Purpose	I Reproduce	II Organize Reorganize	III Invent Generate
1. To learn (metalingual, mathetic)		Summary (B,C) Paraphrasing (A)	
2 To convey emotions, feelings (emotive)		* Narrative/Personal story (A,B)	Open essay (B,C)
3 To inform (referential)		* Letter to an uncle describing a bicycle (A,B) * Self description in letter to pen-pal (A,B) * Formal note to head of school (A,B) * Message to family (A) * Application letter (B,C) * Letter of advice to younger student (B,C) Describing an object (A,B) Describing a process (B,C)	* Reflective essay (B,C)
4 To convince/ persuade (conative)		* Application letter (B,C) * Letter of advice to younger student (B,C)	* Persuasive/ argumentative essay (A,B,C)
5 To entertain, delight, please (poetic)			Open essay (B,C)

Figure 3. The tasks in writing distributed over the boxes in the classification. Compulsory assignments are marked with an asterisk.

The most common of written production at the lower levels of school are based on the students' own experiences. It was therefore natural that one of the first tasks to be chosen should be of a narrative character. Different topics were tested and there was then also an ambition to stimulate to reflections in connection with the relating process. Such topics were for example to try and understand the importance of acting in a certain way, of getting friends, of learning from other people, of failing or being successful or of other events belonging to daily life. The topics were tested separately, but it was decided that in the main investigation the students should be offered to choose from many topics, as it turned out to be a well-established practice to do so in most of the countries taking part.

It was discussed if the students should be allowed to make up their stories or if they should state that what they related was true. A try-out showed that it was impossible to conclude whether a story was based on imagination or not, and impossible to conclude whether a story was based on imagination or not, and for that reason this question was left open.

In English literature on writing the term "prompting" is used about different ways of stimulating the students' appetite for writing. For example one might stimulate their imagination by using the headline "I succeeded - at last!" as a substitute for a headline chosen by the student himself, when the composition is finished and which in a better way indicates the contents of the composition.

Another type of stimulus which, at least for the older students in the nine-year compulsory school, can be more inspiring is to give a finishing sentence which constitutes a conclusion or a statement within a subject field. For the students it is a question of letting the story lead up to this finishing sentence. Also these ways of going about it were compared with each other during the pilot study. It was, however, decided that the more traditional way of setting tasks should be used, even if many students during the study found the method with a finishing sentence amusing and challenging and certain scorers meant that it resulted in better stories.

The tasks that in the following are called 1A, 1B, 1C, 1D and 1E are all different types of messages, and they are directed towards definite and often known readers. Such written tasks often play a remote part in the teaching process. Task 9 also has an informative purpose but like 1E it has the aim of convincing as well as informing. A more detailed discussion of the pragmatic written tasks has been carried on by White and Löfqvist (1988).

Principally at the higher stages of the compulsory primary and secondary school the students get to learn how in speech and writing to argue and practise convincing listeners and readers. Although it is like that in the teaching systems of most countries it was meant within the project that an argumentative task should be tested in all populations.

Attempts were made to find a field of subject that has reference to the students. Earlier experiences showed that it was advisable to attach the task to questions or problems in the student's own home area and to let the student himself word the problem. Through tests it was decided if in the instruction a hint should be given as to what kinds of questions the student should be able to write about, whether the student should be allowed to direct himself to a special reader or not and if the student should have to limit himself to local questions or also be allowed to treat global problems.

The pilot studies showed that it was unsuitable to exemplify a question to write about. Then the students wrote about that only. Also in certain countries it was impossible to find an authority to turn to in order to get a problem solved. On the other hand it seemed to be advisable to turn to a reader of a different opinion. The best result was achieved if the students got almost entire liberty to decide what question they wished to write about.

The reflective composition or essay is the most difficult task. It implies, with a question or a statement as starting-point, trying to create a content using one's own thoughts and reflections. In most countries the students usually get a number of topics to choose from. Therefore the alternatives of giving one or several topics according to the same method as in the case of the personal story were tried out. The try-out showed that it was best to present a large number of topics to the students, but it was also evident what stimuli had an essential effect on productivity and what expressions in the instruction that seemed to be insignificant, when individual topics were tested (Lamb & Purves, 1988). Several topics of a concrete as well as an abstract nature were tried out. Well-known topics for example referred to school, radio and TV, whereas the more general ones treated poverty and loneliness.

Tasks of this kind are usually not given to students at stages corresponding to the nine-year compulsory school. The project leaders were also uncertain about whether to let population B, that is grade 9, do this writing task. But there was a wish not to limit it to population C, which in other countries generally only comprises those students who go on to universities and colleges, thereby indirectly stating that only these students are able to reflect around notions and ideas and to express this in writing. Irrespective of whether a student succeeds or not when working on the task, the intellectual process of trying to presenting one's thoughts in writing is of great importance, and if population B is also included that would emphasize the fact that it is a valuable experience to have an opportunity to work on reflective writing even to students who are completing their compulsory education.

The work implying choice of tasks to cover the domain of writing, the shaping of instructions and stimuli and the testing of preliminary versions has been extensive. Similar efforts have not been accounted for in any Swedish study. In the FRIS project (Martinsson, 1978) the purpose was to find objective linguistic measures and therefore there was an ambition to find a writing situation in which all students wrote about the same topic and where the topic gave stimulus for productivity. Guided by investigations of student interests at the junior stage detailed close-ups of situations were chosen, among others with animals as stimuli, and for these suggestions were offered in the form of headlines.

For his choice of topics Larsson (1984) expressed the demands that 1) the task should lead to writing activity, 2) the interest in the type of task should influence the students' linguistic activities, 3) the task should be usable both in grade 7 of the compulsory school and in grade 2[3] of the upper secondary school and 4) that the achievements of the different forms should be comparable. These demands were made concrete in free writing with pictures as a stimulating basis, and also in an argumentative task.

In the domain of writing the FRIS tasks can best be described as a personal story, whereas the tasks in Larsson's version correspond to "Open essay" and "Persuasive/argumentative essay" in figure 3.

Practice in writing in the Swedish upper secondary school to a large extent has a design which has played a dominating part in planning supplements, teaching aids and also in established practice. Starting in a reading connection with texts, pictures and factual tables, the students have to analyse, work at and pass on views and ideas in topic suggestions principally with an informative or argumentative purpose at the highest intellectual level according to figure 3.

Such writing tasks demand that the students get plenty of time to penetrate the basic material before they start writing. Even if these writing situations have their advantage they are difficult to administrate and too time-consuming in an international study, where the ambition is to test the students on different types of texts in the domain of writing and in a relatively short time. A writing situation designed in accordance with the central tests would probably also feel strange and look complicated to students in other countries without this type of writing practice. Hultman and Westman (1977) question whether the central tests are too difficult for the average student in the upper secondary school. In their summary they say about the task that was the basis of their study that the material had been too extensive and a little too difficult, that the stimulus texts had tied up the student's language and angles of approach and that practically all the students had lacked the critical training that the task takes for granted.

Grundin (1975) does not state his reasons for his choice of tasks for the test in productive writing. The students were asked to express their views of a definite question - How can we make traffic more secure? - and the task was given to all grades in the primary and secondary school with a testing time of 20 minutes.

2.3.2 Writing tasks

The various assignments for writing are reproduced here in the shape they were presented to the students. Som of the tasks were compulsory in each of the participating countries, whereas they were free to use the other if they so wished. In the Swedish study only the compulsory tasks have been used. In the international original they are the following. The writing time has been given in brackets.

TASK 1 A (20 minutes)

Imagine that you visited your uncle who lives in another town. There you saw a fine collection of bicycles in a shop; in the attached pictures you see some examples of these bicycles. The shop has bicycles for boys and girls, many models and colors with many kinds of extra parts (basket, lights, horns). Your uncle wrote a letter promising that he will buy one for you as a birthday present. Choose a bicycle from the illustration. Complete the following letter with a well-organized paragraph describing the bicycle you have chosen in such a way that he will be able to buy the model you really want.

Dear Uncle,

I think it's wonderful to get a bicycle for my birthday! I have thought about the one I would like and now I think I know.

The bicycle I want ____

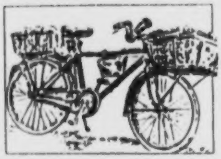
____

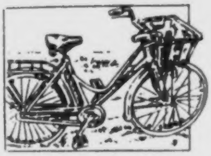
____

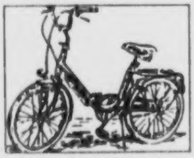
____

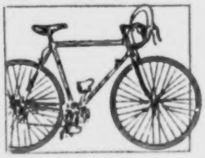
Thank you.

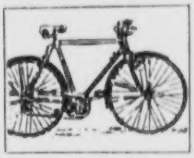
Yours,











TASK 1 B (20 minutes)

You are going to visit somebody (a penpal, a relative) who has never seen you. Write a short well-organized description of yourself (what you look like, what you will be wearing, anything special about yourself, etc.), so he/she can recognize you easily.

TASK 1 C (20 minutes)

Your headmaster has asked you to come to his office at two o'clock. But at one o'clock something unexpected happens to you (an acceptable but not too unpleasant reason) and you cannot go. Write a short message to the headmaster so that he will not wait for you in vain, and apologize.

TASK 1 D (20 minutes)

You arrive home after school to discover that no one is there. You expect that a member of your family will be home soon. You want to go and visit a friend for an hour or two, but you do not want your family to worry about where you are. Write a short message to a member of your family, explaining where you will be, whom you will be with and when you will be home again.

TASK 1 E (20 minutes)

During summer vacation we employ:

-- clerical assistants
-- deliverers
-- park-keepers
-- wrappers

We will interview students who have finished general school. Minimum: One month between June 15th and August 31st. Write, specifying which job you are interested in and when you will be available: City Council, Personnel Department, Arosvar, 9200

You saw this advertisement in the newspaper. You want to apply for one of the jobs. Write a letter asking for an interview.

TASK 5 (60 minutes)

Below are a number of titles for a story. Choose one (1) of them, and write a story that happened or that could have happened to you. You may change the title to match your own experiences if you wish.

Your story will be judged on how good a story it is, on how lively and precise a manner you use to present situations, characters, events, and your own feelings, and on how interesting it is to the reader.

Your story should be at least one page long. When you have finished writing the story do not recopy it, but make any corrections you want to make in parentheses above the words or phrases you change.

1. I Made a Hard Decision
2. I Succeeded - At Last!
3. I Did Something Good
4. My Dream Became Real
5. I Met a New Friend
6. My Most Frightening Experience
7. I Learned to Know How the Other Person Feels

TASK 6 (60 minutes)

Think of something about which you have a strong opinion. The issue could relate to some aspect of your home, school, community or society as a whole.

Write an essay which would persuade somebody who does not share your opinion to change his or her mind and see things your way.

Your composition will be judged on what you have to say, how well you organize and present your point of view and how clearly and effectively you express yourself.

TASK 7 (60 minutes)

Below are a number of observations and questions. Select one (1) and write an essay in which you reflect on what is said and state your own viewpoint. However, in writing your essay you might also take account of different points of view. In your essay you may refer to personal observation, to books that you have read, to films, plays or whatever will contribute to what you want to say.

Your essay will be judged on what you have to say, how clearly you present and illustrate your thoughts and how effectively you express them.

Give your essay a title or use the title given. If you want to correct or change something, you may do it on your original; you do not have to recopy the whole essay.

1. Does watching television make it more difficult to think independently?
2. Many young people today find it difficult to talk to and understand middle-aged people.
3. What might happen if students were to be given more control over what they study and how they study?
4. What might the world be like if the role of women in society really changed?
5. It is preoccupation with possessions, more than anything else, that prevents men from living freely and nobly.
6. Doesn't the heart in the middle of crowds feel frightfully alone? (Charles Lamb)

TASK 9 (20 minutes)

Write a letter of advice to a student two years younger than you who is planning to attend your school and who has asked you to tell how to write a composition that will be considered good by teachers in your school. Write a friendly letter and describe in it at least five specific hints as to what you think teachers in your school find important when they judge compositions.

In each task the students are asked to answer some questions on the last page of the form when they have finished. The questions are identical with the wording of the first one adapted to the task. In the case of 1A they look like this.

Please, answer the questions about the task you have just completed.

1. How familiar are you with the kind of writing task you have just completed: a task which asks you to describe an object?
    - A. Not familiar at all
    - B. Familiar to some extent
    - C. Very familiar
2. How easy or difficult did you find the task?
    - A. Difficult
    - B. Relatively difficult
    - C. Neither easy nor difficult
    - D. Relatively easy
    - E. Easy
3. How interesting did you find the topic (task)?
    - A. Uninteresting
    - B. Relatively uninteresting
    - C. Neither interesting nor uninteresting
    - D. Relatively interesting
    - E. Interesting

2.3.3 The questionnaires

In addition to the writing tasks the students have had to answer two questionnaires. Furthermore, the headmaster and the students' Swedish teacher were responsible for one questionnaire each.

In the headmaster inquiry certain questions are about the municipality where the school is situated, its size, material and cultural standard. Other questions bear reference to the school, its size, general resources and special characteristics.

The teacher questionnaire contains questions on the teacher's educational background and teaching duties as well as on the school and the class that takes part in the study. The latter part is about teaching methods. The teacher has to take his stand about different aims of the teaching of writing and different parts of the subject as well as about different descriptions of preparation, realization and evaluation of essay writing.

In the student questionnaires there are on one hand more general questions about sex, age, educational and vocational aspirations, linguistic background and co-operation in the home between children and adults, on the other questions directed towards reading and writing at home and at school. The latter refer among other things to the student's interest in books and the mass media supply, attitude to writing and possibilities of learning to write.

2.4 Target groups, planning and realization

The study refers to three age groups of students:

> Population A comprises students who have reached or are near the final phase of the class teacher stage, that is grade 6 on the intermediate stage of the Swedish school system.
>
> Population B referes to students at or near the end of the moment at school when a large proportion of them leave school. In Sweden this population corresponds to grade 9 of the senior stage of the compulsory school.
>
> Population C consists of students at or near the end of the upper secondary school. Of Swedish upper secondary school students this means those in the third grade of the 3(4)-year courses and in the second grade of the 2-year courses.

All participating countries are represented by population B. Only a smaller number take part with populations A and C. Sweden is represented by all the three populations.

The students have had to write the tasks that are shown in table 1. Some of them have been rotated in the classes, as each student can only be tested with the use of a couple of tasks. In the table the rotated tasks are underlined. In population A it is thus 1A-D that have been rotated, so that a quarter of the students have written 1A, another quarter 1B etc.

Table 1. The distribution of the written tasks and the number of classes and students

1A	description of bicycle,	5	personal narrative
1B	self-description,	6	argumentative essay
1C	message to the headmaster,	7	reflective essay
1D	message to the home,	9	advisory letter.
1E	answers to advertisements,		

Population	Written tasks			Number of classes	Number of students
A (grade 6)	1A-D	5 6		48	1210
B (grade 9)	1A-C, 1E	5 6 7	9	72	1702
C (grade 2^2, 3^3)	1E, 9	6 7		75	1533
Total				195	4445

The table shows that the nine-year compulsory school students wrote three tasks each and the upper secondary school students two.

The tasks were carried out on different occasions. On some of them the students also had to answer questionnaires. Table 2 shows where the tasks and questionnaires were placed. The writing time was 20 minutes for tasks 1A-1E and 9. For the others the students had 60 minutes at their disposal.

Table 2. The distribution of writing tasks and questionnaires

Point of time	Pop A (grade 6)	Pop B (grade 9)	Pop C (grade2^2, 3^3)
1	1A-D Questionnaire	1A-C, 1E Questionnaire	1E, 9 Questionnaire
2	5	5-7	6-7
3	6	9 Questionnaire	Questionnaire
4	Questionnaire		

The tables show that students at different age levels have had to write the same tasks. 1D however is exclusively meant for grade 6. All have had to write task 6 but otherwise two grades have been linked together. The purpose of this is to make comparisons possible of the achievements of students at different stages.

The number of compositions on the different topics is shown in table 3.

Table 3. Number of compositions per task

Task		Grade 6	Grade 9	Grades 2^2, 3^3
1A	description of bicycle	297	416	
1B	self-description	296	411	
1C	message to the headmaster	301	420	
1D	message to home	299		
1E	answer to advertisements		406	719
5	personal narrative	1187	572	
6	argumentative essay	1187	565	717
7	reflective essay		488	676
9	advisory letter		1606	732
Total		3567	4884	2844

3 Schools, teachers and pupils in Sweden 1985

The following descriptions of schools, teachers and students are chiefly based on a selection of the answers in the questionnaires described in 2.3.3. They principally refer to conditions that have no direct connection with the teaching of Swedish. Such conditions will be treated in the next chapter.

3.1 Schools

In this short description of the schools principally such conditions are described as are not known by means of public statistics. They chiefly refer to the headmasters' evaluations of the material and cultural standard of the school municipality and the school. A more extensive description can be read in a separate report (Löfqvist, 1989f).

3.1.1 Material standard

The local education allowances are considered by 3/4 of the headmasters in the compulsory school to be average in comparison with other similar municipalities. The headmasters in the upper secondary schools are of a more differentiated opinion. Half of them mean that their own municipality gives about the same economic support as others, but 1/4 consider their own municipality more generous, and 1/4 meaner. It is difficult to have a definite opinion about why the upper secondary shool headmasters differ in their views from the headmasters in the compulsory schools. Possibly the reason is that their schools are principally located in larger population centres.

There are pupil libraries at all schools, and practically all of them have equipment to show TV and video programmes as well as films. Reading rooms and study halls which the students may use during lunch breaks or free periods exist at three schools out of four.

For teachers there are at 2/3 of the compulsory schools a library with occupationally-oriented literature. This information should be understood to imply the schools have a set of such professional educational literature as can usually be found in the staff rooms. It is not very common to have a special room exclusively intended to be a teacher library more than possibly at upper secondary schools.

The headmasters answered the questionnaire during the school year 1984/85, and as early as that almost all senior levels of the compulsory school and upper secondary schools owned computers. The number per school probably differed - from only a few to complete sets for regular teaching. This information is now out of date but it shows how fast the fitting out of computors has developed at the higher levels, not least thanks to special State grants. On the other hand only 1/4 of the intermediate-level schools had access to computers.

During the 1980s complaints have been made by among others schools, students and parents that the economy measures concerning teaching aids have been driven too far. It has been said that the schoolbooks are often too old, that they are torn and scribbled all over and that the students are not allowed to have books of their own.

In most schools the students may use their books when working at home. But there are senior-level schools where no students have the right to make use of schoolbooks when doing their homework.

In a number of schools at all levels certain groups of students have no opportunities of working at home using their schoolbooks. This refers only to a small number of compulsory schools, but in the upper secondary school it concerns two out of ten. The answers do not indicate if the headmasters aim at specific type of education.

3.1.2 Cultural standard

The headmasters have estimated the travelling distance from the school to different cultural resources. Within close reach of nearly all compulsory schools and upper secondary schools there are public libraries, places of worship, recreation premises and sports centres. Furthermore it is possible within two hours and from a dominating number of the schools irrespective of stage to reach a bookshop, a cinema, theatre, art gallery or museum. Only a small number of compulsory schools can be said to be isolated from such a cultural supply.

It seems to be easy to get to places with upper secondary and adult education. Even college education can be reached after a few hours' journey, although the number of educational programmes at the centre of higher learning is in many cases limited.

In addition to the existence of a cultural supply the headteachers have also been asked to estimate the cultural standard of their own municipality compared with other similar areas of the same size, with the same geographic position and the same economic conditions. Irrespective of stage the headteachers think that the cultural standard in their municipalities is average or above average. Only a small number think that it is below average.

The headteachers do not however estimate equally much the inhabitants' participation in the cultural activities. They rather seem to think that people become engaged about as much as in other municipalities.

3.1.3 Internal work

In connection with the latest revision of the curriculum for the nine-year compulsory school greater responsibility was put on individual schools to design on their own aims and ambitions in a work schedule. Such a plan should indicate concrete aims and purposes structured in stages among other things for the basic skills of speaking, reading, writing and arithmetic. For the nine-year compulsory school there is a duty to show in a work schedule the way education has been planned, whereas such local curriculum work has been carried out only as an experiment in the upper secondary school.

Local aims of the teaching of writing have been designed in 3/4 of the intermediate-stage schools but only in half of the senior-level schools. In upper secondary schools it seldom happens that there are local work schedules for writing according to the headmasters. There are however centrally drawn up supplements for the planning of writing to be used by the teachers.

That not all nine-year compulsory schools have made up plans for the teaching of writing in 1984/85 may be due to the fact that other parts of the work scheduled have been given priority to. It may be that the larger number of positive answers from the intermediate-stage schools are due to the fact that there Swedish is a dominating topic with a larger number of lessons.

The compulsory school headmasters are satisfied with the way the parents support schoolwork. As might be expected they are more satisfied with the parents' contributions at the intermediate-stage than at the senior-stage. In the upper secondary schools there is a decrease in the parents' commitment and those who consider the support of the parents unsatisfactory are more numerous than at the senior-stage.

The number of pupils' clubs at the schools is small. Only on very rare occasions in the upper secondary school are there unions for literature, Swedish writing, foreign languages and mathematics. At about 1/4 of the schools students can come together in their leisure time to be engaged in theatre, art, music or science. Such club activities are however unusual in the nine-year compulsory school.

Further it seems to be a fact that in certain schools there is a stronger club tradition than in others. They can have many clubs of different sorts, whereas others may have only a sports association. Such an association exists in three schools out of four at the intermediate-stage and in most senior-stage and upper secondary schools. Especially in upper secondary schools they have political and religious associations, and more recently organizations for computers, ecology and films have been created.

In many schools the students edit their own papers. At the intermediate- stage they work together to issue such a paper in 1/5 of the schools and in the upper secondary schools almost half of the schools have a pupils' magazine of their own.

3.2 Teachers

The description of the teachers' background and working conditions is given here in a short summary. A more detailed account with a comparison with the teachers' conditions in some other countries in the study has been shown in a separate report (Löfqvist, 1988c).

3.2.1 Age and education

The teaching of Swedish in the nine-year compulsory school and in the upper secondary school is carried out by women in two cases out of three. The female dominance decreases with higher stages.

The average age of teachers of Swedish at the senior level and in the upper secondary school is about 40. The spreading is the widest in the upper secondary school, where every fourth teacher is over 50 and every tenth under 30. The intermediate-stage students met younger teachers. Nearly half of them are between 31 and 40 and every tenth under 30.

Nearly all teachers have a theoretic and pedagogic education. Among the teachers at the senior stage, however, 13% state that they lack formal education in Swedish.

The teachers seem to mean that their education about the writing of compositions has been more directed towards the treatment of the finished composition that towards the preparatory work and the writing process.

The subject teachers seem on the whole to commit themselves more than class teachers in activities related to Swedish. Even if their active participation in pedagogic activities both during school hours and outside them is relatively limited, it seems as if topic specialization stimulates their interest. It is probable that special knowledge and proficiencies within the topic means better opportunities of getting commissions as consultant, lecturer, expert etc.

3.2.2 Working conditions

The intermediate-stage teachers almost exclusively teach one class only, whereas senior-stage and upper secondary school teachers have several classes. Many teachers, especially at the senior-level schools and in the upper secondary schools, have reduced teaching obligations. The senior-stage teachers on average teach more subjects, have more different courses but the same number of classes as the upper secondary school teachers.

Preparative and corrective work in Swedish is of course different for class teachers and subject teachers. There is an increase with each higher grade partly because work gets more time-consuming the older the students are, partly because the senior-stage teachers and especially the class teachers have more subjects to prepare and control. The intermediate-stage teachers say that more than half of them daily use one hour or less on this work, about half of the senior-stage teachers use two hours, whereas the upper secondary-stage teachers think that they use about 3 1/2 hours. As a large proportion of the teachers that answered the questionnaire have reduced teaching duties, it probably follows that the homework of those doing full-time teaching is more extensive than the average values.

3.3 Students

In two reports (Löfqvist, 1989b; 1989d) an extensive description is given of what has here been summarized about the students' social and linguistic background, their reading interests, TV viewing and attitudes.

3.3.1 Academic and vocational ambitions

The students in the compulsory school want to go on studying. Their propensity for study is high. Their hopes of future jobs can be compared with the present distribution of these occupations in the gainfully emplyoed population between 25 and 64 years of age. A distinct tendency seems to be that the students hope to get opportunities to work within occupational groups with a high status. Especially occupations such as officials at high level or a medium level are attractive. The number of students who aim at occupations within such groups is larger than the percentage groups of the total population.

3.3.2 Immigrant students

The percentage groups of parents who use another language than Swedish as their mother tongue is larger for intermediate-stage and senior-stage students than for upper secondary school students. They make up 18-21%. Parents who have another mother tongue and have their children in grades 2^2, 3^3 are fewer or 14%. In each grade about 10% of the students have learnt to speak another language than Swedish. These students should in all likelihood find it more difficult to do themselves justice when practising written work in Swedish than immigrant students who were born in Sweden and have Swedish as their first language. How well immigrant students have learned to use Swedish in speech and writing depends on at what age they started learning Swedish. The predominant part of them state that they started learning Swedish before they were ten.

In all age groups immigrant students always or nearly always use their first language at home. But they do not read newspapers in their home language. When they speak to others, read or write Swedish is the common language. There is however a greater tendency among them to use the home language in conversations with fellow-countrymen of their own age and grown-up friends and for letter writing, which is probably due to the fact that the students when outside school are together with fellow-countrymen in most cases.

3.3.3 Supply of books and papers

The students have been asked to estimate the number of books in their homes. Such estimates tend to be a little vague and approximate, and the question was also worded in such a way that the student was aware that the answer was not meant to be exact. The approximate quantitative estimate that the alternatives imply, however, are sufficient to give an idea of the students' possibilities of finding literature at home.

In most homes there are a lot of books. Only a small proportion of the students say that the number is smaller than 50. One third of them consider that in their homes there are about a hundred books but half of them think that they have considerably more.

In nearly all homes there is at least one daily paper. Four out of ten students have an opportunity of reading at least two daily papers. Still more students can read at least two weeklies or monthly magazines. In more than half of the homes there are two or more such publications. On the other hand 15% of the homes do not seem to buy any weeklies at all.

3.3.4 Interest in reading

Among the students of the compulsory school those who often visit libraries are the intermediate-stage ones. About 2/3 of them go there several times a month, and only a small proportion never find their way to a library. At the senior-stage schools the inclination to borrow books decreases strikingly, and three out of four students in grade 9 visit a library only once a year. The interest in the books in the library is not very much greater as regards the upper secondary school students. The number of visits to libraries, however, is not the measure of an interest in reading. Even if people do not go to a library very often they may just as well like reading.

Chiefly the students in grade 6 read comic papers. Two out of three students spend about 2-4 hours a week on such reading. Every fifth student says that he studies comics at least one hour per day. With increasing age the interest diminishes and half of the upper secondary school pupils do not read comics at all. For the others it is usually a question of not more than a few hours' reading per week altogether. The reading of comics as regards the senior-stage students is of a scope between the consumptions of the other student groups.

The difference that can be found between pupils at different levels concerning the reading of comics does not exist for other types of light reading. It is indeed astonishing that they all show almost the same values.

On average the students spend about 3-4 hours a week on reading other things than comics. If we add other kinds of reading to this - partly comics, partly more school oriented reading which however does not imply homework - the average values are about 8 hours a week for the higher stages and well over 9 hours for the intermediate stage due to the fact that those students read more comics. Together with the time the students read during the lessons and during their homework reading takes up a considerable part of the day.

3.3.5 Homework

The pupils at the intermediate stage state that they do homework every day. It takes about 1/2 - 1 1/2 hours for most of them to do their homework. At the senior stage about the same amount of time is devoted to homework. But almost every tenth pupil says that he does not do homework at all. Also at the upper secondary school an equally large proportion say that they do not spend time doing homework which might be due to the different purposes and aims of the educational programmes. The students of the theoretic educational programmes must devote more time to homework than those who are vocationally-oriented. The upper secondary school students have a heavier work-load totally. 70% have at least one hour's homework every day and

25% work for two hours or more.

3.3.6 TV and radio

In addition to the time devoted to reading and homework the students watch television and listen to the radio every day. The pupils of the intermediate-stage seem to be the greatest consumers of TV programmes. On average they watch TV about two hours during a schoolday. Every tenth pupil does not watch TV at all but on the other hand just as many sit for four hours or more in front of the TV set. The difference between the senior level and the intermediate-stage students is insignificant. Half of the upper secondary school students however say that they devote one hour at the most to watching TV. On the whole they use less time watching TV than the other students even those days when they do not go to school. The average value for the TV watching of the upper secondary school students during schooldays is about 1 1/2 hours and during holidays about two hours. The average values for the others are during schooldays about two hours but during holidays well over three hours for grade 6 and just under three hours for grade 9. Half of the intermediate-stage pupils usually watch TV at least four hours during holidays.

The information that the students have given about the time they spend on doing homework, reading, watching TV and listening to spoken radio programmes may give an idea of how they on average use the time after school. The values are of course approximate but the following table can give a rough estimate of how much time the students spend on the different activities each day.

Table 4. Approximate time per day for different activities outside schooltime for pupils in grades 6, 9 and 2^2, 3^3

	Doing homework	Reading	Watching TV	Listening to the radio
Grade				
6	about 1 hour	about 1 1/2 hours	about 2 hours	well over 1/2 hour
9	about 1 hour	well over 1 hour	about 2 hours	barely 1 hour
2^2, 3^3	well over 1 hour	well over 1 hour	about 1 1/2 hours	barely 1 hour

3.3.7 Family climate

The students have been asked to estimate how often it happens that children and grown-ups in their families do different things together.

There is considerable accord of opinion among the pupils. The family members go together helping each other in the home and they have their meals together to the same extent in all student groups. Almost every day in all homes they eat together, often help each other with different household chores but they seldom seem to work together within any organization. The fellowship in duties and activities of this kind are not dependent on the student's age. It is therefore natural that the average values for the different age groups should be almost the same. It is only concerning homework that there is a considerable difference. The pupils in grade 6 get help a couple of times per week, whereas the upper secondary school students almost completely manage on their own, which perhaps is often due to the fact that grown-ups do not possess knowledge enough to be able to help.

In activities that can be said to be built on a common interest and which are performed for fun the younger pupils have higher average values than the others. They more often go together with adults than the teenagers who do not often go with their families to different gatherings, the cinema, exhibitions etc. For all these descriptions there is a difference between grade 6 and the others. But the upper secondary school students more often discuss family affairs with grown-ups. Also it seems to be more usual that they take part in debates at home on for example ethics, religion and politics. This difference may be due to the students' greater maturity and interest in such matters.

3.3.8 Students' world of ideas

The description of the students can be completed with an account of what might be called their world of ideas. An analysis of content of their compositions on task 6, according to which they were to write about something that they have a very definite opinion of, has given the following information.

About 1/3 of all students irrespecitive of grade write about school. Garme (1988) also found a strong interest among the pupils in writing about school.

The pupils in grade 6 write about different school problems with an emphasis on questions of school rooms, school meals service and school transport. In grade 9 their interest is more directed towards teaching and especially in grade 3[3] the students discuss scoring and teaching problems. Even if many are critical and want to have changes brought about, most of them on all age stages are positive to school.

An equally large proportion or 1/3 of all pupils write on national social problems. In grade 6 the pupils write most of all about sports, hobbies or animals that they favour or back up. In comparison with the others the choices of subject of the intermediate-stage pupils are concentrated on a smaller number of fields. The pupils in grade 9 share with grade 6 an interest in sports and hobbies but they discuss just like the upper secondary school students most of all relations between people.

The last third of all pupils write about their homes, about their own society or about world-wide social life questions. They are distributed among the three items into three groups almost equal in number. The compulsory school pupils have views on the rules at home that were set up by their parents and want to change them. The upper secondary school students above all discuss equality questions. Almost all of them mean that they ought to help with household matters. Most of them are satisfied with or rather content with the society they are living in. Their dissatisfaction is most of all attached to the lack of spare-time occupations. The global questions deal with war and peace, nuclear weapons and underdeveloped countries. Several pupils in the compulsory school, especially in grade 6, have written about these subjects. Only a few upper secondary school students have chosen such subjects.

In their choice of subject the pupils seem to be influenced by their surroundings, and naturally also by the teaching they get at school and information from the mass media. The difference between the grades is often naturally conditioned by age and maturity.

4 The teaching of writing

A description of how the teaching of writing is organized in the classrooms from the pupils' point of view has been given by me in the report "Students write compositions" (1988b). How the teachers mean that it is carried out I have given an account of in "Teachers of Swedish" (1988c). Furthermore I have made a comparison between the teaching of writing in curricula and in practice (1989a). The content of these reports is the basis of the following chapter.

4.1 Aims and principal items

The majority of the teachers state that they attach great importance to the aims of the teaching of writing. The development of imagination and emotions is especially at the intermediate level considered as more essential than communicative skill and intellectual language practice. But the teachers at the higher stages mean that it is more important that the pupils learn how to use their own language correctly.

The teachers do not think that the curricula are as significant to what they include in their daily teaching as their own estimates of what the pupils will find useful when they leave school. The content of teaching is guided in comparison with schoolbooks or central tests also principally by what the teachers themselves consider to be important for the pupils to learn.

The teaching of Swedish in grade 6 is chiefly characterized by linguistic usage and by linguistic structure exercises. The pupils spend much time on orthography. In grade 9 teaching is dominated by the reading of fiction and literary history, but grammar and the history of language also get an essential part of the time. The usage and structure of language seem to play an obscure part in the upper secondary school. The reading of literature and the history of literature are the most important items.

The skill of writing compositions is considered important by teachers of all stages.

4.2 Writing practice and preparations for writing

With increasing age the pupils get more and more teaching of the different phases of writing. They increasingly learn to gather material, to give this material structure but also during the writing process to keep their plan open for additions and changes. Not least important is for the pupils to practice the mastering of the formal skills. The reader should not be disturbed by grammatical and orthographical blurs and mistakes.

Exercises in writing technique are used chiefly at the intermediate stage but also in grade 9. Especially orthography and punctuation are often practiced. But those skills are trained less often in the upper secondary school, where the pupils usually master them. Also exercises to practice linguistic variation are used more often in the compulsory school. In the upper secondary school more emphasis is put on practice in how to use and cite source material and also to give a composition a systematic and logical structure.

Analysing different kinds of text is a natural feature in writing practice. The teachers at the intermediate stage generally use the compositions of their classes as a basis of discussion. At the senior stage and in the upper secondary school it is more common to use texts written by professional authors. Central tests from earlier years are used as practice material especially in the upper secondary school, where tests of that kind are compulsory.

The pupils at the intermediate stage get more practice in doing different kinds of tasks than their older schoolmates. This applies to practice in writing technique, linguistic variation, in the use of different types of text but not however practice in how to deal with the content. Such practice is carried out chiefly by the upper secondary school teachers. That the intermediate-level pupils practice more than the others is probably due to the fact that these pupils need more formal training. The number of lessons in Swedish is also considerably higher in grade 6 than in the other grades. The teachers in those classes can thus spend more time on the teaching of writing.

Before the pupils start writing a composition the teachers can discuss the topic with them and help them with ideas and suggestions. They can also give support as to choice of words and phrases. A working method like that is above all used at the intermediate stage. At other grades it only happens occasionaly that the teachers give such extensive help. The usual thing seems to be that the pupils write about unprepared topics.

4.3 Writing tasks

Both teachers and pupils have mentioned what writing tasks are used most of all at school, and the pupils have also described the tasks that they prefer to work on.

In the following tables the writing tasks in the first two columns are ranked according to means in a three-point scale. In the third the order is built on the percentage points each subject or type of subject has received.

Table 5. What the pupils in grade 6 are asked to write and what they want to write

1 What the pupils say that they are often asked to write		2 What the teachers say that they often use		3 What the pupils like best to write	
Fantasy essay	2.89	Story about one's own experience	2.60	Fantasy essay	23%
Summary	2.45	Story of an event	2.44	Story about one's own experience	8%
Story about one's own experince	2.35	Summary	2.42	Composition about nature (animals)	8%
Review	2.20	Short story	2.34	Adventure story	8%
Description of object	2.06	Private letter	2.19	Factual writing	4%
Conversation	1.98	Book review	2.13	Detective story	4%

Giving one's imagination a free rein and writing stories about one's own experience is what the pupils do in most cases in grade 6. This is also what they like best to write about. We can see that this type of free imaginative writing (Fantasy essay) can be found in column 1, in column 2 (Story of an event, Short story) and in column 3 (Fantasy essay, Adventure story, Detective story). They are both allowed to write about what they themselves have experienced and they like to do it.

Writing summaries of the content of a book or a story as well as reviews of books occur quite often according to both pupils and teachers, but the pupils are not fond of devoting themselves to such tasks. Nor are they eager to write descriptions or dialogues, which they say often occur at school. Some of them, however, have favoured factual writing, above all topics dealing with nature.

Table 6. What pupils in grade 9 are asked to write and what they want to write

1		2		3	
What the pupils say that they are often asked to write		What the teachers say that they often use		What the pupils like best to write	
Fantasy essay	2.23	Story of an event	2.29	Fantasy essay	17%
Summary	2.10	Story about one's own experience	2.29	Story about one's own experience	10%
Description of an event	2.06	Short story	2.22	Love story	8%
Story about one's own experince	1.97	Personal view	2.20	Description of hobby	6%
Review	1.92	Book review	2.18	Adventure story	4%
Conversation	1.56	Summary	2.13	Composition about nature (animals)	4%

The majority of the pupils in grade 9 prefer to make up their own stories, often about love and with an element of excitement. They also quite often get an opportunity of writing free topics in that sense. Pupils and teachers agree that personal experiences often constitute material for compositions.

As in the case of grade 6 pupils and teachers agree that books often form the basis of summaries or reviews. But the pupils do not state that they wish to write such compositions. Some pupils, however, like to write about their hobby or about animals and nature.

Table 7. What pupils in grades 2^2, 3^3 are asked to write and what they want to write

1 What the pupils say that they are often asked to write		2 What the teachers say that they often use		3 What the pupils like best to write	
Summary	2.34	Personal view	2.29	Fantasy essay	17%
Review	2.22	Summary	2.28	Story about one's own experience	10%
Fantasy essay	2.00	Book review	2.17	Description of problem	7%
Story about one's own experience	1.94	Story about one's own experience	1.98	Factual writing	5%
Personal view	1.90	Short story	1.90	Description of hobby	5%
Reflections	1.66	Story about an experience	1.87	Love story	4%

As has earlier been pointed out also several upper secondary school students like to write their own stories on self-chosen topics and only few students are negative. These tasks occur in the teaching, but not very often. It is more common to write summaries and reviews, which the students have not mentioned among subjects that they like to write about.

4.4 Realization and scoring

The totally dominating working method seems to be that each pupil writes his composition on his own. Only on isolated occasions is the composition allowed to be the result of joint work, usually at the intermediate-stage. Especially the younger pupils can get help during the writing process, but in most cases the pupils get no feedback until the composition has been corrected.

Most of the teachers read and comment on all that the pupils write. At lower stages though shorter products may be evaluated in another way than through the teacher's examination, for example by means of feedback from other pupils or teachers.

The usual method at each stage is that the pupils get their compositions supplied with either corrections or markings from which they can make a revision themselves. It often happens that the intermediate-stage teachers discuss the compositions individually with the pupils, and one such individual discussion is something that the other teachers try to arrange once during the term. These two methods can naturally be used in combination. After the teacher has corrected the compositions, the content and the corrections may be discussed individually with the pupils. But mistakes that many pupils have made can be talked about with whole class. This is obviously something that happens more often in the upper secondary school than at other levels.

5 Compositions and scoring

5.1 Method of scoring - a theoretic and empiric background

In nearly every Swedish investigation on writing at school evaluation is discussed. The most essential problem is to find a suitable measure of quality when the pupils' achievements are evaluated.

Different factors might influence the pupils' performances. Some of them can be attached to the writers and the writing occasion. The pupils may for example be inspired for writing on certain days but not feel motivated at all on other occasions. The time of the day can play a part just like experiences before the time of writing. Even if the pupils should get the same topic to write at different occasions, one may suppose that the results might differ.

Other circumstances can be attributed to the subject. The pupils' capacity can vary due to the type of task, stimuli and additional advice and instructions.

In the IEA study each pupil is however asked to write three tasks of different kinds - in the upper secondary school two tasks - which to some extent ought to neutralize such an influence, which might be dependent on the writer, the occasion or the subject. The attention has, however, chiefly been directed towards the rater. Coffman (1971) has summed up the problems which can briefly be expressed in the following way: different raters tend to evaluate the same performances differently, individual raters tend to estimate the same performance differently on different occasions and these differences tend to increase, if the pupils' freedom as regards the written task gets bigger.

Thorén (1951) pointed out lack of agreement among experienced teachers' estimates of upper secondary school compositions. The concord among the teachers was weak not only in their scoring of the same compositions. Even the teachers' verbal comments on the language of the same composition was seen to change from "expressive" and "lively" to "poor" and "thin".

An increased reliability in the evaluation of a composition can be reached, if many raters get an opportunity of evaluating it, and if the mean of their estimates may make up the final marks of the composition. The more raters that are resorted to, the better the possibility of reaching the "true" value. Björnsson (1960) showed that it is enough to have three raters to attain a correlation of at least 0,90 with the "true" mark, which according to his investigation was the mean of well over a hundred researchers. The FRIS project (Lindell, 1980) used three raters both in the teacher evaluation, and in the so-called creativity estimate. Larsson (1984) has with direct reference to Björnsson made use of three raters. The project "Swedish measures" (Hersvall et al, 1974) asked both teachers and journalists to evaluate compositions and got the assistance of three from each category. At the standardization of the written standard tests for the compulsory school in grade 9 the compositions are evaluated by five teachers (Eneskär, 1988).

One possibility of diminishing subjectivity has been multiple evaluaion, that is an evaluation made by two or preferably three raters and with the mean of their estimates as a measure of the written achievement. Different variations in English and American investigations are reported by Purves, Gorman and Takala (1988) in their survey of the development of the evaluation process. The best solution is the same as the one Björnsson found. This multiple evaluation combined with an analytic marking of the written tasks has characterized the IEA study.

An analytic marking means that several evaluations are made of some essay. A number of aspects are decided on and then there is an ambition to try and concentrate the evaluation on these. It is possible, for example, to put special emphasis on the content of the composition and first evaluate only this. At the evaluation of the content it must not matter how the material has been arranged, that is the outline of the composition, and not the linguistic quality either. Besides content, outline and perhaps especially linguistic style are aspect that are evaluated separately.

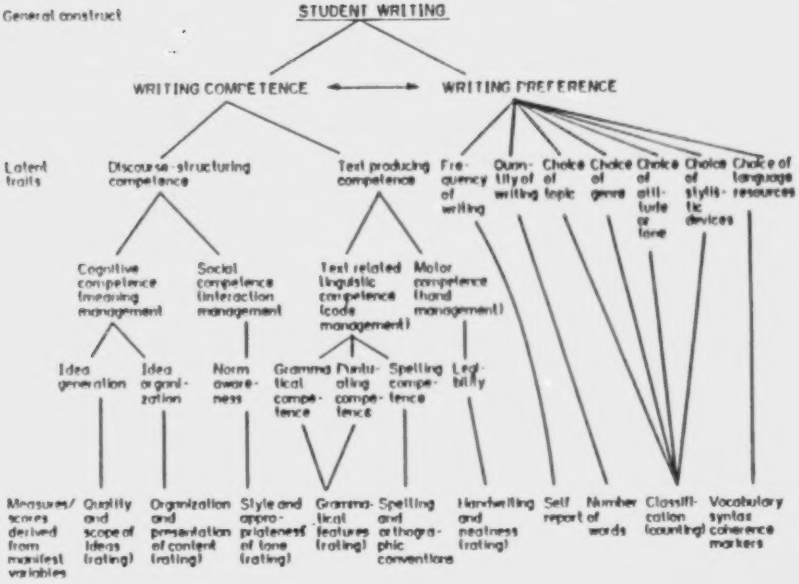
The method which is the most common in Swedish schools is a general impression marking. In such an evaluation there is probably a rather vague opinion of what different parts create totality and what importance they are considered to have. Thus it is possible to take into consideration content and outline, linguistic variation and correctness and to try and put them together to create a totality. It is probable that the main impression of a composition in most cases is based on a more or less deliberate merger of different evaluation aspects.

Even if the teachers in Sweden in most cases use a general impression marking, many teachers work hard to modulate this with comments on a special aspect, for example the content of the composition. The evaluation of totality at higher grades is most often given in the form of marks, whereas shorter or longer comments may make the pupil observant of for example content and outline. Quite often comments are given in the margin, even if it is more common to use that for linguistic correction marks.

Both the FRIS project (Martinsson, 1978) and the project "Swedish measures" (Hersvall et al, 1974) have discussed the possibility of profiling the evaluations so that they refer to different aspects. Both projects refer to Page and Paulus (1968) who made experiments with profiled evaluation. Each composition in their investigation was marked as to the aspects ideas, organization, style, correctness and creativity. The different aspects, however, turned out to receive very similar estimates. The intercorrelation was .80 and more except for the aspect of correctness, which correlated .65 up to .69 with the aspects of content, organization and creativity. In both the Swedish projects it was decided to use marks from the total evaluation when the analyses were carried out.

Larsson (1984) compares total evaluation with an analytic marking and finds with reference to Pages and Paulus that an analytic method is far too time-consuming without being superior in objectivity. Larsson uses a total score shaped like a mark.

In the discussion within the IEA project the concept of writing in school was looked at from two starting-points. It is possible to talk partly of writing skill or writing competence, partly of writing preference. By the last term is meant what attitude the pupil has towards writing, how often he writes, what topics he prefers to choose, what genres etc. At the evaluation, however, writing competence is of principal interest.


General construct
STUDENT WRITING
WRITING COMPETENCE
WRITING PREFERENCE
Latent traits
Discourse-structuring competence
Text producing competence
Frequency of writing
Quantity of writing
Choice of topic
Choice of genre
Choice of attitude or tone
Choice of stylistic devices
Choice of language resources
Cognitive competence (meaning management)
Social competence (interaction management)
Text related linguistic competence (code management)
Motor competence (hand management)
Idea generation
Idea organization
Norm awareness
Grammatical competence
Punctuating competence
Spelling competence
Legibility
Measures/ scores derived from manifest variables
Quality and scope of ideas (rating)
Organization and presentation of content (rating)
Style and appropriateness of tone (rating)
Grammatical features (rating)
Spelling and orthographic conventions
Handwriting and neatness (rating)
Self report
Number of words
Classification (counting)
Vocabulary syntax coherence markers

Figur 4. A general construct of student writing (from Purves, Gorman and Takala, 1988)

What then does writing competence imply? The discussion led up to the construction in figure 4, and on this the evaluation system was based.

No matter if one has to express oneself in speech or in writing, one must have competence to give structure to what one wishes to say or write. One must be able to find ideas and clues, judge one's ideas and distinguish what is essential from what is not, use one's knowledge and make comparisons, be creative and imaginative, that is what in figure 2 is briefly called idea generation. In this cognitive competence, however, there is also the capability of organizing one's thoughts. They must be linked together in such a way that the readers can assimilate the account. We usually talk about organizing the material in larger or smaller units. This cognitive competence is a condition for the writer's success in getting the readers to understand what he wants to say. He may for example wish to inform or instruct the readers, increase their knowledge within a certain area and then analyze and make connections clear. In order to make the writer's message reach the readers the content has to cover the subject and have a logical outline.

The writer may however have another purpose than only to inform. The intention may be to influence the readers, to make them change their opinions or attitudes or perhaps also to create a preparedness to act in a certain way. If one wants to make the readers feel or think in a special way one must indeed know something about the readers and have what in figure 4 is called norm awareness. One must know about the rules, norms and attitudes that have reference to the readers and guide their way of thinking. In working life it is essential for example to advertising people and journalists that they are biased in the wording of their message if they intend to reach and influence their readers.

Even at school there are demands on the pupils that they should adjust what they write to their readers. Letters for example must be outlined with the addressee in mind. Style and tone ought naturally also to change with regards to the topic. To be aware of what guides the reader is here called social competence. One has to be conscious of present norms and adjust one's text to the receiver.

What we have now discussed here can be made concrete in three aspects of evaluation. The first one refers to the content of the composition with its wealth of views and ideas and their quality. The second one has reference to the capability of planning and presenting a content. The third estimates the ability to find a suitable style and linguistic tone going with the content.

Writing skill also includes competence to produce a text. One has to be able to express one's thoughts in words. One essential part is the linguistic competence to give a correct form to what one wishes to say. One must have acquired a skill in expressing oneself according to present grammatical conventions. In addition to this one has to follow orthographic agreements.

The competence to produce texts also implies motor skill in writing down one's thoughts by hand. At school the pupils generally use script or they print and only as an exception do they use a typewriter. There are therefore demands on them to design the letters well, so that the reader can easily understand the text without being prevented by obscurities from seeing what the text says.

The competence to produce a text can also be divided into three aspects of evaluation. The first one refers to knowledge of present grammar usage and power to put generally accepted writing rules and language norms into practice. The second refers to spelling, writing in two words or in one word, the use of capital letters, that is what we usually summarize as orthographic skill. Last of all an evaluation can be made of the legibility of the text, when the clarity of handwriting is especially emphasized.

Within the project the participating countries agreed to use a method which combines a total estimate with an analytic evaluation. All of them were to use the three aspects of competence to structure the account. Other aspects could vary with each country. The model that we used for the evaluation of the Swedish compositions is described in detail in section 5.2.

5.2 The scorers' guide

For the guidance of the scorers a guide had been worked out within the international secretariat. It needed a national adaption and for Sweden's part it has been made in consultation with the members of the national committee. Together with the guide there is a supplement with examples of students' compositions to illustrate what is being said.

In the guide the total evaluation is called Overall impression. The aspects that have been discussed in the previous section have the shorter terms Content, Disposition and Style. What in the schedule is called Grammatical and Orthographical conventions have in Sweden been put together under the term Readability. Handwriting and neatness correspond to Legibility.

The evaluation is made by means of a five points scale, where 5 is the highest mark. For Legibility only the marks 5, 3 and 1 are used. A special scoring sheet had been worked out.

The guide first gives a short description of the aspects and practical instructions for the evaluation. Then it is indicated what should be considered at the evaluation of the different aspects. On the evaluation of *Overall impression* the following instructions are given.

> "Read the composition in a normal fashion, without stopping to correct, mark, underline segments of the composition or to write comments. Record your first impression

> of the composition as a whole. Indicate on a scale from 1 (inadequate) to 5 (excellent) what you think of the composition as a whole. Do not change this score once you have recorded it.
>
> The evaluation of the composition as a whole should consider all possible aspects, by means of which a composition can be estimated, for example content, disposition, linguistic correctness, orthography, style, linguistic tone. Lower marks must therefore not be given only because for example the number of spelling mistakes is large or because the planning of the material is defective. Nor should a lower total evaluation be made if the pupils have chosen to write in a genre other than that expected. It is a well-known fact that the pupil's handwriting may conduce to the composition being overestimated or underestimated respectively. Therefore observe that poor handwriting must not be considered as a drawback to a performance which is in other respects good.
>
> Do not try to adjust or in any other way adapt the proportions of the number of compositions to different marks, for example by making use of a distribution according to the normal curve. The compositions that you evaluate are only part of a large number. Your feeling of how insufficient or excellent the composition is should only be related to the international scale, that is the so called national benchmark composition.
>
> Avoid as much as possible to reread the composition to make your total evaluation or to change the marks you have given from the beginning guided by your first reading. In most cases it is not necessary to spend more than a few minutes on a composition.
>
> You must not use half marks and you are not allowed to write more than one mark, for example 2/3 or 4/5."

After the overall impression had been registered in the scoring sheet the partial aspects were estimated. It is emphasized in the instructions that each aspect should be evaluated separately and that the rater should not try to make his gradings the same for different aspects. Nor must an average of the estimates of the partial aspects correspond to the overall impression.

For *Content* the following instruction was given.

> "The aspect implies supply of relevant ideas and views in the composition. During your evaluation you should consider the character of the task. Your estimate is also based on somewhat different factors depending on the task, as it is difficult to define in a general way the demands on content. A summing-up of the demands on each task is given in the instruction for the evaluation of the separate tasks."

IEA International Study of Written Composition
The Department of Educational Research

Scoring sheet

Scorer

School number																				
Pupil number																				
Task number																				
Overall impression (1-5)																				
Content (1 - 5)																				
Disposition (1 - 5)																				
Style (1 - 5)																				
Readability (1 - 5)																				
Legibility (1 - 5)																				
The task cannot be evaluated																				

Task numbers

11 Description of bicycle
12 self-description
13 message to the headmaster
14 message to the family
15 answer to advertisement

51 difficult decision
52 managed
53 good deed
54 my dream
55 new friend
56 worst experience
57 fellow-creature

60 personal opinion
71 watching TV
72 understand old people
73 decide at school
74 woman's role
75 desire for possessions
76 loneliness
90 advice on writing

If the task cannot be evaluated the following should be taken down
7 misunderstanding
8 illegible
9 blank paper

The instructions related to the special topics implied for example in the case of task 1 characteristics for the bicycle to be identified, in the case of task 2 sufficiently observable and distinctive characteristics for a identification, in the case of task 3 essential facts that ought to be mentioned in the message.

The aspect of *Disposition* just like Content had an overall instruction and detailed instructions going with the tasks.

> "This aspect concerns the way in which the written material as a whole has been organized. The structure of the compositions can vary considerably. The evaluation should bear in mind to what extent there is a logical consistency in the discussions and a coherence among and within the individual paragraphs."

During the evaluation of the compositions it turned out at a preliminary investigation that it was difficult to mark off the aspect of *Style*. Above all it seemed as if linguistic errors in the compositions tended to influence the estimate. Therefore great care was devoted to an exemplification and a description of the aspect.

> "In this connection style indicates the pupil's choice of appropriate language with regard to an imaginary reader or addressee, the topic under discussion and purpose of the task. In marking a judgement about the writer's control of stylistic features you should bear in mind the options available, especially in terms of vocabulary and sentence structure.
>
> The pupils are expected to observe certain linguistic norms suitable for the linguistic tone that is used and for the organization of the text, for example in a letter with an application for a job. An evaluation of this skill will be made in this aspect. If however the pupils do not state name, address or other important facts in an application for a job, this must be referred to the aspect of content.
>
> When the pupils write to certain people, for example the headmaster of the school, they must use a suitable linguistic tone. There are situations in which one should not write in such a familiar tone that it may be understood as being impolite or impudent. In a letter to somebody of the same age however a very familiar tone could be the natural thing. One problem is that there can be different opinions on where to find a limit between what is very colloquial, informal or normal written language.
>
> If the pupils use different linguistic means of expression to keep the reader interested, it should be referred to style and tone.
>
> It should be put down to the pupil's credit if he succeeds to make his teacher absorbed in what he has written. Humor and wit should for example be attributed to

style. Some common and characteristic deviations from suitable style and linguistic tone are illustrated in the following:

Lack of variation in combinations of clauses and sentences

That pupils can be very uncertain about stylistic convention is especially obvious when they have to try and link together a number of sentences. There are several examples showing that too often they use "and", "then" and "when". In writing such a style is considered as immature.

Example: In the morning mother woke me up. First I ate and then I cycled to school. When I came home I did my homework. Then I would ride away and play handboll. When I came home I ate and watched TV.

Repeated use of words or expressions

In spoken language words and expressions are often repeated. Such repetitions may give the impression of monotony when they occur in writing.

Example: When I was out walking in the wood I walked by a large oak which was in the middle of the wood. When I got past the oak I heard a buzz. I stopped and walked back. When I was standing near the oak I saw a big hole underneath.

Choice of words and phrases

When we talk we use different little words that imply attitudes and points of view. We sometimes double certain little words or combine the sentences in a vague way. In a more experienced written language one usually clears away such little words. Another kind of uncertainty concerns the choice of words.

It can be a little distorted and inexact, without the reader for that reason missing the meaning.

Example: I decided to go and find out if they had already gone. I was so tired that I took the lift. When I was on the 19th floor the lift stopped.
My friend had to be in hospital for a week then when came home, he was happy.
It is not only racists that are stupid but also those who for example do not like Jews."

Even for Style there are special instructions attached to each task. These emphasize in many cases that the scorers should consider if the pupil has chosen suitable words, phrases and larger linguistic units with consideration for the reader, for example in letters to an older relative, to the headmaster or to an authority.

As can be seen in the instruction linguistic errors are not included in Style. What is generally considered as unacceptable deviations from linguistic rules and norms are classified as *Readability*. In the Swedish evaluation system there has been made a union within this aspect of what in the schedule of writing competences has been divided into grammatical and orthographic conventions. The instructions read as follows.

> "Imperfections concerning orthographical and grammatical conventions disturb the reader's understanding of the content and make the text difficult to comprehend if there are serious linguistic errors. You are to estimate to what extent the pupil knows such conventions.
>
> The deviations from the convention of the written language and orthography concern spelling, the writing of words in parts or as one-word units, the use of capital or small letter and punctuation. If the pupil does not follow grammatical conventions the meaning of the written text becomes obscure and vague, as the grammatical relation between words, phrases, sentences and larger units is not satisfactory.

Poor and indistinct punctuation

The written language aims at stating in a perspicuous way the relation between sentences and clauses. In the spoken language such a relation is indicated by means of intonation and the use of pauses. In the written language different punctuation marks are used. Faulty or lacking punctuation, which results in breaches of grammatical convention, is generally called faulty sentence structure. Some common types of errors are the following.

A No distinction between independent sentences.

Example: It started with my mother and father going out for a walk I and my brother were alone there it was in the night and had turned dark suddenly the telly went out. Today we were to move to our summer cottage who we are it is me, my mother, father and my two brothers and sisters.

B No distinction of demarcation between clauses within a sentence.

> Example: When we were sitting in the boat I said how many years old are you? I am twelve that is what I am too I said. We talked for a long time and we took a swim. And had great fun. You yes he said will you have picnic with us yes I will.

Coherence

When the coherence of a text is incomplete readability is made more difficult. In different ways deficiencies in coherence between parts of sentences or between sentences that follow after each other influence understanding.

Of frequent occurrence are for example error of reference, omission of words, use of incorrect verb forms, incorrect choice of conjunction.

> Example: I asked where she came from and that I came from Sweden.
>
> I think that smoking is awful. First of all they contain much nicotine and tar.
>
> As Maria is very obstinate, therefore she did not give up. Then I tried to think of something that might convince her.
>
> Then suddenly her machine broke down said she also that the machine was bad.
>
> It costs a tremendous amount of money and when you smoke others who do not smoke may get smoke into them. Besides it might begin burning if you do not put it out farefully and also you must realize that matches and lighters will also cost a lot of money.
>
> When it was my turn to be interrogated he suspected that I had taken it, when he saw that I had two, but then my pal said that I had taken her wallet because she had dropped and I had taken up and was going to give it to her, but just then the man shouted that he had lost his wallet and forgotten to give it to her."

There were no special instructions for the aspect of Readability concerning the different tasks. The general instructions for both this and the following aspects were considered sufficient. For *Legibility* the following instruction was given.

> "The evaluation in this aspect is meant to distinguish between compositions in which the handwriting is clear, regular and easy to read and those where the handwriting makes reading more difficult and also obstructs the understanding of the written text.

The following conditions ought to be observed, as they influence legibility:

1) the design of individual letters

2) the total impression of the handwriting, for example the uniformity of the inclination of the letters

3) spacing, that is adequate distance between lines and between words

The evaluation is carried out in three grades

Mark 5	clear, distinct, accessible on first reading
Mark 3	some pauses required in reading, some reliance on context to distinguish particular words
Mark 1	sentences illegible on first reading. Heavy reliance on context to interpret the writer's meaning. Only few of the words at once legible."

5.3 Evaluation results

At the beginning of this chapter the discussions within the IEA project leading up to the method of evaluation have been described. While reasoning about writing competence that pupils need to structure their text, we distinguished three aspects which might be evaluated separately. Ideas and views were summarized as content, ability to organize content as disposition and capability of finding a suitable tone of language as style. Writing competence however also implies ability to produce text. The evaluative aspects were made concrete partly in a linguistic aspect of being able to apply grammatical and orthographical conventions, readability, partly in a motor aspect of being able to shape the text legibly and tidily, legibility.

The compositions have been evaluated in these five aspects by means of a five-grade scale, except for the last-mentioned, where only three marks have been used. In addition to this each composition has been given an overall impression mark. A guide with instructions for the scorers has also been presented earlier.

Different sets of internationally evaluated compositions have been used in the investigation. To the scoring guide have been added a number of compositions with the purpose of making the instructions more concrete. Besides there has been an internationally scaled set of so called "benchmark compositions" for each task, by means of which a scorer would be able to adapt his evaluation level to the international norm. To make it possible to revise the national evaluation according to the international scale the scorers have had to mark a number of compositions in what was called "the calibration set", an internationally scaled set of compositions. These and the benchmark compositions have been used for different calculations, which are accounted for in this chapter.

An evaluation method ought to be designed in such a way that a composition is scored exactly the same by different raters to make it at all possible to make sensible and correct comparisons of pupils' achievements. The method described in previous sections with special instructions and with particular compositions as a frame of reference partly for an overall evaluation, partly for an evaluation of limited aspects has been intended to constitute such an instrument. If this functions in a satisfactory way the requirements to make comparisons of pupils' achievements in different countries are fulfilled. As an evaluation of compositions contains strong subjective ingredients, there is however a risk of national estimates deviating from the international norm. Should this happen, there ought to be a possibility of correction. The calibration compositions mentioned earlier were meant to be used to state whether a deviation had taken place and in addition to that to serve as a guide for corrections.

In this chapter we will discuss questions on the correlation between national and international evaluation. Two comparisons must be made. It is essential to check that national scorers keep to the same scoring level and also that they route the compositions in the same order as is the case in international scoring. One section therefore deals with average values of marking and another with the correlations with international scales.

A number of compositions, the so-called "repetition essays", have partly been evaluated by all the raters, partly been scored by the same person a second time after a certain interval. There is therefore a possibility of determining both the correlation between the raters and the constancy of each individual rater.

These different calculations constitute a check-up of the evaluation. After this description we will show the results of the main investigation giving the marks of the written tasks, the correlations between the pupils' marks of different written tasks as well as for some topics, the correlations between the marks in the different aspects in which the compositions received marks.

The pupils got three questions after each written task. The first question meant that the pupils should indicate the topic's frequency during their lessons using one of the answers Not at all, Rather common or Common. The second referred to the degree of difficulty on a five-grade scale from Difficult to Easy, and in the third the pupil stated in five steps, how stimulating the subject was from Boring to Great fun. The questions are in the following called attitude questions, even though the first one strictly speaking does not concern attitude. Calculations have been made on the correlations between the attitudes and between marks and attitudes.

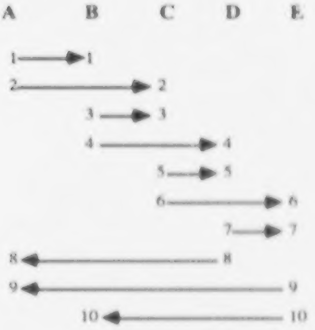
Each composition has been given marks by at least two raters. Their average mark values have been estimated and also the correlation between them. These calculations have been made on the main study.

First however we will describe the model we have used to make comparisons between several markers possible.

5.3.1 Model of scoring

The designing of the model of scoring was made with a view to one composition being evaluated by two raters. The compositions of all tasks except number 6 have been scored by two independent raters with a pair wise connection according to the table below. The raters are indicated by means of letters and the compositions by means of figures. Within the group of five raters comparisons can be made.

Raters



Each figure in the table indicates a bundle of compositions, e.g. 50. In each bundle there may be compositions from different stages. The essays have been marked with an identification number, but neither through this nor in any other way can it be concluded what grade the composition belongs to if it is not revealed by the content. Rater A starts with bundles 1 and 2 and after evaluation gives them to B and C, rater B starts with bundles 3 and 4 and then send them on to C and D etc. All the raters have in this case eventually scored four bundles, that is 200 compositions.

Among the compositions in all the bundles have been distributed the calibration and repetition essays that were described earlier. The repetition essays, which were 20, had been copied. All the raters thus received the same essays. Rater A thus had 10 repetition essays in bundle 1 and 10 in bundle 2. He received them after some time for a repeated evaluation in bundle 8 and 9 from D and E. The calibration compositions were distributed in the same way with the difference that in some of the tasks they did not get a double evaluation.

As raters some 50 teachers with an experience of teaching Swedish at all the three stages were consulted. Each group of five raters consisted of teachers from the stages that had written the task.

Task 6, the argumentative essay, has been written by all three populations. Each essay has been evaluated three times and therefore the above model has not been possible to use. Instead groups of raters with three in each have been set up. In each of those groups have participated one teacher from the intermediate stage, one from the senior stage and one from the upper secondary school stage.

5.3.2 Training and control of scorers

Training of scorers

The training of the scorers aimed at making them accustomed to the international scale. In connection with the training sessions calculations were made of their agreement with the international evaluation. After the training period similar calculations were made, when the compositions in the main investigation were estimated.

Before work on a task started, the raters were invited to information on the IEA project. The guide that was described earlier was discussed and the instructions were studied and illustrated by the essays in the appendix. After that each of them was told to work at home on the scoring of the compositions that constitute the benchmark set. The raters however did not then get to know what marks those compositions had received during the international evaluation. They put down their estimates using work sheets that were sent in form processing. Because these compositions were used as practice material the raters were forced to work on them actively. The compositions were to serve as their frame of reference from then on. Calculations were made of their means. Furthermore the correlations with the international scoring of these compositions were decided on, as well as among the raters themselves.

At the next meeting the results were discussed. Each rater received his own means and correlation coefficients. Each composition came up for discussion. Motivations were given and discussed for extreme marks concerning compositions with a wide spread. Comparisons were made with the guide and its collection of examples. The raters were told to consider deviations in the means and during their continued work make a careful correction. At the same time they were warned of the risk of a compensation that might easily get too strong and lead up to an opposite deviation.

For further practice at the meeting the raters received a number of compositions to work on. These were chosen at random from the investigation and had not been marked before. The individual estimates of the raters were taken down on the board for each composition and aspect. A discussion was started, if the differences between the raters were considerable. There was an ambition to arrive at an agreement and concordance in their evaluation.

It was, however, also important that after the practice session the same calculations should be made as during practice. As each rater was to score the calibration essays which were wedged in among the Swedish compositions, it was possible once more to make comparisons with the international scale. Thus the mean score value was calculated for each rater as well as the correlation with the international scale and between raters.

Still one more calculation was made. For all tasks, except number 6, 20 compositions were chosen for a repeated estimate. These were distributed at random within the material. The correlation between the first and the second scoring of these repetition essays was caluculated for each rater, and this therefore constitutes a measure of constancy in the individual rater's scoring.

All compositions have been studied in original, except the calibration and repetition essays. As these were to be scored by all, they have been photocopied. In that way they have differed from the others by appearance. The raters were informed that certain compositions would appear twice and they were told not to memorize especially how they had scored such compositions.

The following calculations have thus been made of the teachers' scoring before and after practice.

Before practice	**After practice**
Comparisons of the raters' mean scores with the international scale (the benchmark set).	Comprisons of the raters' mean scores with the international scale (the calibration set).
Correlation between raters and the international scale (the benchmark set).	Correlation between raters and the international scale (the calibration set).
Correlation between raters (the benchmark set).	Correlation beween raters (the calibration set toghether with the repetition essays).
	Constancy among raters (the repetition essays together with, for some tasks, the calibration set).

Mean scores

If marks for compositions are to be equivalent, the scorers must demand equally much when evaluating the pupils' performances. They have to apply the same scale. Then their means will be identical. To make it possible to compare the performances of pupils in different countries, the evaluations in these countries must be carried out according to the same scale. To make sure that the raters follow the scale, their evaluation has to be checked.

As has earlier been described the raters had to score the essays in the benchmark set before practice. After practice when evaluating the composition in the investigation they were asked to score calibration essays for each task. By comparing the means of the raters for both sets of essays with the international values one can determine to what extent they have managed to follow the international norm. In the diagram below it is possible to make such comparisons for each task.

Furthest to the left tasks have been given and within brackets the populations that wrote the task. Tasks 1A (the description of a bicycle) and 1B (the self-description) are registered twice, because grade 6 and grade 9 have been estimated by means of separate scales. In task 1C (the message to the headmaster) however, the two grades have been estimated according to the same scale.

The number of raters for the tasks has varied. For some with a smaller number of compositions there have been 5 raters, for the others between 10 and 27. For each task the mean value of the raters' mean scores before and after training has been calculated. This value has been compared with the international mean score of the task. Thus for example the international mean of the essays in task 5 was 2.91 in the benchmark set. The raters' scores for these compositions before practice gave on average a lower mean value. It was 2.62. In the diagram the difference of -.29 between these values has been stated. The international mean for the calibration essays was 2.28 and the raters' mean value 2.24. It was therefore almost the same value. The difference of -.04 after practice has been indicated in the diagram.

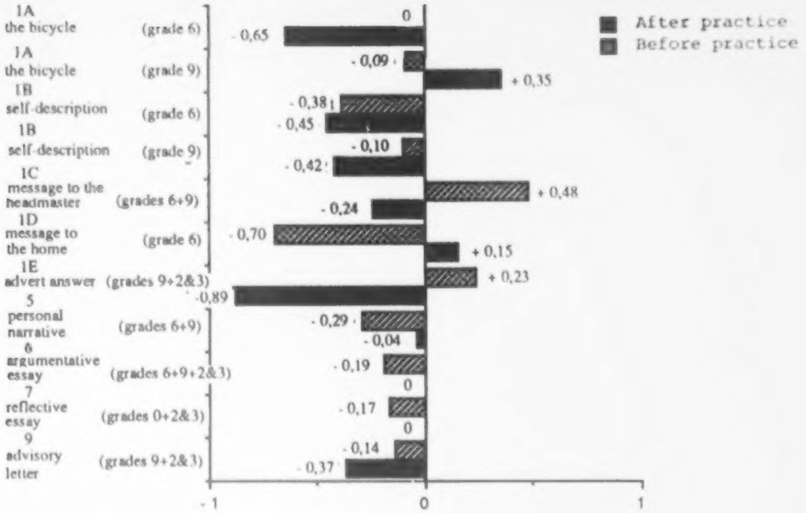


Diagram 1. The raters' mean scores compared with the international scale before and after the practice session.

The expected development was that a difference between the raters' means and the international evaluation would be reduced by means of the practice session. The differences after training ought to be smaller than before. That is so in about half the cases. But otherwise the deviations have increased and also changed direction. Thus the raters were above the international scale in task 1C before practice but at a lower value after practice. The deviations from the international norm is for task 1E almost one step on the marking scale. As this difference refers to the average of 20 raters' values, it indicates that the deviation in the case of several individual raters is considerably greater. Why are the changes not the ones that were expected?

The reasons for the deviations could be many. In some cases it is probable that the raters have adjusted their scoring too heavily. If their mean values in the benchmark set were too high or too low, an adjustment might lead to overcompensation. In task 1C the raters were .48 steps on the scoring scale above the international mean on the fist occasion, but they were .24 below on the second occasion. It might be assumed that such compensation may also have influenced other tasks.

The translation of the compositions may have played an important part. The compositions in the benchmark set, which were Swedish compositions exclusively, and those in the calibration set, which came from countries that were not English-speaking, had been translated into English before the international evaluation. In a special investigation of the importance of the translation to changes in the quality of the compositions we have been able to verify that translation results in improvements of the language (Löfqvist, 1989c). Even compositions that have correct linguistic form improve. What, on the other hand, can be characterized as impairment of freshness and originality by the translations does not probably result in a lower marking, at least not to the same extent. It is therefore to be expected that a set of compositions of varying quality by means of translation can on average be better than the originals. When the compositions in the benchmark set were used as practice material for the Swedish raters, copies of the Swedish original versions were used. It is therefore quite possible that these circumstances can help to explain that the Swedish raters are below the international value before practice in the case of most of the tasks.

As the raters are connected with one of the stages in their teaching it might be possible that they favoured or treated unfairly one stage or other when they made their evaluation. Their personal attitude and experience might result in especially lenient or strict evaluations of pupils' performances at other stages. A comparison between the teacher groups' mean scores for the same tasks, however, do not indicate any such difference according to what is shown i diagram 2, which has been designed in the same way as the previous one. In the tasks where the difference is greatest it amounts to about .1 step on the scoring scale except in the reflective essay, where teachers at the upper secondary school and the senior stage of the compulsory school differ after practice. The upper secondary school teachers are above and the senior stage teachers below the mean value of the international scale.

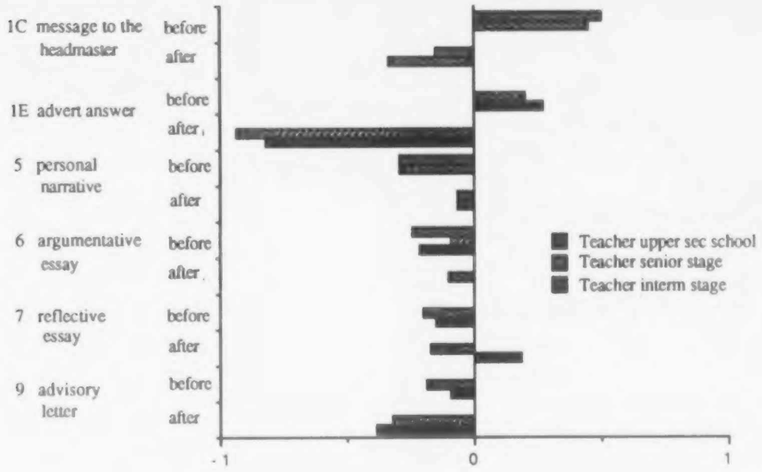


Diagram 2. The raters' means per teacher group compared with the international scale before and after the training session.

By way of explanation one might suppose that the teachers changed their demands during their evaluation. While working on several hundred compositions they might perhaps get more lenient or more strict in their estimates. A measurement of such a change can be indicated by the evaluation of the repetition compositions. A register of the means from the first and the second evaluations is given in table 8. In task 6 there were no repetition compositions.

Table 8. Means at the first and second evaluations of the same composition

Task		Average marks	
		1	2
1A	description of bicycle (grade 6)	2.8	2.9
1A	description of bicycle (grade 9)	2.4	2.5
1B	self-description (grade 6)	2.9	2.8
1B	self-description (grade 9)	2.7	2.7
1C	message to the headmaster	2.6	2.5
1D	message to the home	3.0	3.1
1E	advertisment answer	2.5	2.4
5	personal narrative	2.9	2.8
7	reflective essay	2.8	2.9
9	advisory letter	2.9	2.9

The values in table 8 are the means of all the raters within one task. The difference between the first and the second evaluations is at the most .1 step. In the case of some individual teachers it was greater but never more than .3 step. These values indicate that the teachers are so constant in their evaluation that the deviations from the international scale cannot reasonably be due to the teachers without any grounds changing their degree of demand.

An important reason may very well be the insufficient number of compositions included in the two sets. The benchmark set consisted of 10 up to 12 essays and the calibration essays were on average only 7 to 8. The teachers' deviations in their scoring of a small number of compositions can result in considerable effects in the mean values.

The most essential reason, however, is likely to be found in the difficulties of constructing internationally equivalent patterns. The compilation of different sets of compositions was made by groups of experts with members from different countries. The experts marked the compositions independent of each other. When they were of divergent opinions a discussion in the group could lead to an agreement on the marks. It would take too long to give a detailed report here on the process but in short an analysis of the procedure leads up to the coclusion that the different groups of experts demonstrated a systematic difference in the evaluations. One of them turned out to be too lenient, another was too strict. This lack of agreement means that a national benchmark set cannot constitute such a frame of reference for evaluation as had been planned. Nor can the calibration essays be a satisfactory instrument to make international comparisons possible. No definite demands of agreement with an international scale can then be made on national raters.

Correlation with the international scale

In the previous section there was a comparison of the raters' ability to keep to the international scale level by means of a study of the mean scores. But even if these are similar, a further measure is needed to show how good the agreement is between different estimates. One has to know whether the raters rank the compositions in the same way. This concordance is expressed in table 9 in correlation coefficients. The values for each task constitute means for all the raters' correlations within one task. In the left-hand column we have the values for the correlation with the benchmark set, in the right-hand column for the correlation with the calibration set.

Table 9. The raters' correlation with the international scale before and after the practice session.

Task			Before training	After training
1A	description of bicycle	(grade 6)	.81	.94
1A	description of bicycle	(grade 9)	.81	.77
1B	self-description	(grade 6)	.74	.77
1B	self-description	(grade 9)	.88	.87
1C	message to the headmaster		.62	.77
1D	message to the home		.79	.71
1E	advertisment answer		.87	.78
5	personal narrative		.83	.88
6	argumentative essay		.84	.87
7	reflective essay		.71	.83
9	advisory letter		.79	.86
Mean value			**.74**	**.80**

The correlations are with one or two exceptions high and rise after the practice session. The mean value for the correlations before the training session is .74 and after the training session .80.

A high correlation only implies that the scorers have followed the international norm to such an extent that they have placed the compositions in the same order. To be able to state if the agreement is good totally one must compare with the table of means. It shows that in the case of tasks 5, 6 and 7 the international and national means agree with each other and in addition to that the correlations are high. In the light of what was said in the previous section about imperfections in the international scale it is however hazardous to express one's opinion on what the agreement between the Swedish and the international evaluation altogether really signifies.

Correlation between the scorers

The compositions in the benchmark set for a task were evaluated by all the scorers. They have also all scored the calibration essays and the repetition essays. Therefore it is possible to calculate the correlation between the teachers' scoring both before and after the training session. The last-mentioned calculations are based on 25 up to 30 compositions within each task. The figures in table 10 are the mean values for all the raters within one task.

As has been stated earlier the number of raters per task varies. Som of the practically oriented tasks have been evaluated by 5 teachers, whereas the others have usually been evaluated by at least 10 teachers.

Table 10. The correlation between raters before and after the practice session.

Task		Before training	After training
1A	description of bicycle (grade 6)	.70	.83
1A	description of bicycle (grade 9)	.77	.73
1B	self-description (grade 6)	.51	.73
1B	self-description (grade 9)	.86	.75
1C	message to the headmaster	.74	.72
1D	message to the home	.73	.79
1E	advertisment answer	.87	.77
5	personal narrative	.77	.76
6	argumentative essay	.82	.78
7	reflective essay	.79	.70
9	advisory letter	.71	.74
Mean value		**.75**	**.75**

The correlations between the raters are strong. One would have expected the correlations to have been higher after the training session just like it was at the comparison with the international scale according to table 9, but this has not occured. The mean value is .75 both before and after practice.

As has already been mentioned the correlations are high both as regards the correlation with the international scale and between the raters. They are higher than is generally found in comparable investigations (Hersvall et al, 1974; Martinsson, 1978; Eneskär 1988).

One explanation of the higher values in this investigation is probably the choice of compositions. The compositions which are among the benchmark set and are all written by Swedish students, had been selected as "typically" Swedish ones, threes and fives. They were to get a final scoring by an international group of raters. This process has been described more in detail in an earlier report (Löfqvist, 1988a). The marks for the compositions have in part been changed in the international scoring, but most of them have been allowed to keep the marks they represented according to some Swedish teachers.

The compositions chosen for calibration come from different countries. For a compositions to be accepted for calibration it was claimed that all the international raters were agreed on the mark. There was a high demand on these compositions that they should be "typical" of the mark they were considered to represent.

That these two groups of compositions can be said to be "typical" of their mark level means that they have a character of their own. It is a fact that a five on which all the raters have agreed is not at all typical but rather extreme. This distinctive character of certain calibration compositions and also compositions in the benchmark set make it easier for the raters to arrive at the same mark in this limited group of compositions.

In the group of essays which has been the basis of calculations of the correlations after the training session also the repetition compositions were included. These have been chosen at random out of the material and can therefore not be considered as "typical" in the same sense as the others. In combination with the calibration essays as a basis of calculation it is, however, probably that they result in a stronger correlation than could be expected.

One can suppose that it would be easier for the raters to come to an agreement, if they have to rank compositions that have been written by students from different age groups than if the students belong to the same group. For example it ought to be easier to distinguish performances in task 5 that were written by both grade 6 and grade 9, than in task 1A which, it is true, were written both by grade 6 and grade 9 but which have been evaluated separately. Especially in task 6, which was written by all populations and where the essays have been scored according to the same scale, the differences between the best and the weakest performance ought to be conspicuous. The correlations, however, do not show such a difference. They are altogether high both before and after the practice session.

A measure of the differences between the teachers from different stages gives the correlation for teachers inside the stage compared with the correlation of teachers outside the stage. A comparison of the correlation between the teachers within one stage can therefore be made with the agreement with the teachers outside that stage. These values have been calculated for some of the tasks.

Table 11. The correlation between raters within and outside stages

Task	Stage	Correlation between teachers within a stage	Correlation with teachers outside that stage
1C message to the headmaster	M	.72	.72
	S	.73	.72
5 personal narrative	M	.76	.76
	S	.77	.76
6 argumentative essay	M	.86	.80
	S	.88	.84
	U	.77	.80
7 reflective essay	S	.65	.70
	U	.72	.70

The correlation within a group of teachers does not in any of the cases seem to be considerably stronger than the correlation with the teachers outside. There is thus no reason to suppose that the raters' experience of a certain stage would have an influence on the scoring and favour or be unfair to other groups o students.

Constancy of scorers

It is well-known that the scoring of a composition is dependent on exterior factors without any connection with the quality of the composition. A rater is exposed to an irrelevant influence that can be difficult to neutralize. A composition can for example receive a higher estimate if it is evaluated after a number of weak performances, but it may be less appreciated if it follows after compositions that have received higher marks. Changes in the evaluation can also be due to tiredness, for example if a large number of compositions are estimated in a row, the students' handwriting may have an influence on the evaluation etc.

The possibilities of reducing such effects are limited. It has, however, been possible to certify if the outcome of evaluations that have been made during a shorter period differs from results when plenty of time has been available. For certain tasks the teachers got ample time and were able to distribute the evaluation of about 300 compositions in a way that they found suitable. They were asked to limit the number of compositions that they evaluated on each occasion. Other tasks were evaluated during a week's time with about 125 compositions per day.

What can be evaluated are the marks that the teachers have given the repetition essays. They were 20 for each task and had been distributed at random among the other compositions.

For each scorer the correlation has been calculated between the first and second evaluation of the repetition essays. The means for all the raters within each task are given in the following table. The task marked with * have been scored during a short period of time. In these tasks also the calibration essays were scored twice, so that the correlations have been calculated for 25 up to 30 compositions.

Table 12. The correlation between the first and second evaluations of the same compositions

Task				Correlation
*	1A	description of bicycle	(grade 6)	.90
*	1A	description of bicycle	(grade 9)	.84
*	1B	self-description	(grade 6)	.90
*	1B	self-description	(grade 9)	.88
	1C	message to the headmaster		.81
*	1D	message to the home		.88
*	1E	advertisment answer		.87
	5	personal narrative		.80
	7	reflective essay		.78
*	9	advisory letter		.83

The constancy is high. Furthermore and contrary to expectation it is highest for those tasks that were scored during a short period of time with long evaluation sessions. The result can be compared with what was indicated in the section on the teachers' means for the repetition essays. The difference was in no case greater than .1 step on the scoring scale between the teachers' first and second evaluation of the same compositions.

To sum up there is support for the statement that those about fifty teachers from different stages that as a group scored the compositions show an unchanged marking. The group does not materially change their scoring level and on the whole keep the marks once given. In the case of individual teachers, however, evaluation may vary somewhat from one occasion to another.

5.3.3 Result of the main study

Scores for the writing tasks

In this section we will first state the number of compositions that were evaluated and those that for various reasons did not receive any marks. In addition to that the means for the partial marks in the different aspects within different topics and grades have been given. The means are related to the international scoring scale. After that the students' marking profiles are discussed, that is their performances within the different types of tasks. The partial marks have been correlated with each other within a number of tasks in each grade to examine to what extent the raters vary in their scoring of different aspects of a composition.

Number of compositions and their distribution among topics

Each rater according to the instructions was to give one overall impression mark and five partial marks to each composition or possibly instead indicate if the topic had been misunderstood, if the composition was illegible or if the pupil had refrained from writing. With two raters for a composition not only the marks will vary. There are also differences of opinion between raters whether a topic should be considered as misundertood or if the text ought to be regarded as illegible. Thus it might happen that one of the raters has meant that a student has misunderstood the topic whereas another has felt that the composition was worth getting a mark. A composition may also have received a mark from one rater whereas another has indicated that the student's handwriting was illegible.

In table 13 and the two that follow the values from the first evaluation of the compositions have been given.

Table 13. Number of compositions in grade 6 per topic

Topic		With marks	Misunderstanding	Illegible	Blank	Total
1A	description of bicycel	288	9			297
1B	self-description	267	29			296
1C	message to the headmaster	292	9			301
1D	message to the home	295	4			299
5	personal narrative	1149	33	2	3	1187
6	argumentative essay	1071	100	4	12	1187
Total		3362	184	6	15	3567

At the first scoring of the composition in grade 6 in all 184 were considered as misunderstanding compared with 169 at the second scoring. There were 6 compositions in the first evaluation and 3 in the second that were regarded as illegible. The differences are slight not only as regards grade 6 but also concerning the other grades, which are accounted for later, so it is possible to refer to the result of the first evaluation during a discussion of the relations between compositions that have been given marks or have not been suitable for grading.

Task 6 was difficult for grade 6 to write as has earlier been pointed out. 12 pupils have not been able to write anything at all, and several have told about their school without arguing, they have only written a sentence or so or else they have related an experience. These compositions have been described as misunderstandings. Proportionally more students, however, have misunderstood task 1B, the self-description.

In grade 9 several have refrained from writing as is shown in table 14, but in many cases this can probably be interpreted as a protest against taking part in the investigation and may not be due to incapacity. The number of misunderstandings is the largest in task 7, which above all is due to the fact that the students have misinterpreted the instruction. They have tried in one and the same composition to express opinions on all the topics. Some students have only written a few sentences, and these efforts have also been regarded by the raters as misunderstandings.

Table 14. Number of compositions in grade 9 per topic

Topic		With marks	Misunder-standing	Illegible	Blank	Total
1A	description of bicycle	410	-	1	5	416
1B	self-description	403	6	1	1	411
1C	message to the headmaster	409	10	-	1	420
1E	advertisment answer	402	3	-	1	406
5	personal narrative	544	20	1	7	572
6	argumentative essay	526	33	-	6	565
7	reflective essay	402	73	-	13	488
9	advisory letter	1555	30	1	20	1606
Total		4651	175	4	54	4884

The number of misunderstandings is largest in grade 6 and diminishes with each higher grade. It is first of all due to the fact that certain of the intermediate stage pupils have found it difficult to understand how to write task 6, but also because the older pupils generally have more practice in and knowledge about how a written task should be handled. The number of misunderstandings i lowest among the upper secondary school students as can be seen in table 15.

Table 15. Number of compositions in grade 2^2, 3^3 per topic

Topic		With marks	Misunder-standing	Illegible	Blank	Total
1E	advertisment answer	709	1	1	8	719
6	argumentative essay	666	29	-	22	717
7	reflective essay	631	30	1	14	676
9	advisory letter	715	7	1	9	732
Total		2721	67	3	53	2844

The share of blank answers, however, was largest in grades 2^2 and 3^3 but lowest in grade 6. The higher share of the upper secondary school students is in most cases due to the fact that they have not been willing to take part in the investigation. From several answers it can be concluded that they have not felt motivated to write, as the investigation of the school had been planned to take place during some of the last lessons of the spring term.

In tasks 5 and 7 the pupils were to choose among different topics. But it was possible for them to alter the headline, so as to make it better suited to what they told about. The first of the tasks meant that they had to tell about a personal experience, and the other involved reflecting on a topic and expressing one's own thoughts.

The complete wording of the topics is to be found in chapter 2. In table 16 they have been compressed into a number of keywords. The raters had to decide what topic the pupil might be supposed to have written about in those cases when the composition had no headline or if it had been changed by the pupil. The raters disagreed only on some exceptional occasions. In the table the opinion of the first rater has always been given. Blank forms, illegible compositions and misunderstandings which could not possibly be attributed to a separate topic are accounted for on lines 5:0 and 7:0.

Table 16. Tasks 5 and 7 distributed among partial topics in grades 6, 9 och 2^2, 3^3

	Grade 6 Absolute numbers	%	Grade 9 Absolute numbers	%	Grades 2^2, 3^3 Absolute numbers	%
5:0	20	2	20	3		
5:1 hard decision	44	4	48	8		
5:2 succeeded	160	14	83	15		
5:3 good deed	27	2	25	4		
5:4 dream became real	342	29	117	20		
5:5 new friend	254	21	106	19		
5:6 frightening experience	323	27	150	26		
5:7 how other feels	17	1	23	4		
7:0			73	15	42	6
7:1 TV watching			100	20	202	30
7:2 young people - older people			94	19	148	22
7:3 pupils decide			116	24	124	18
7:4 woman's role			21	4	49	7
7:5 desire for property			29	6	47	7
7:6 loneliness			55	11	64	9

It has been pointed out before that the pupils in the compulsory school prefer to write about such matters as they have been allowed to make up themselves. Those subjects - "My dream became real" and "My most frightening experience" - have also tempted the majority of the pupils. The instruction allows the pupils not to tell their own experiences, which most of them have not done either. The compositions, especially in grade 6, are imaginative adventure stories. Their personal experiences seem to be used much more in their stories about when they made a difficult decision or performed a good deed.

The parallelism between grade 6 and grade 9 as to choice of topic is obvious. One might discover that grade 9 is a little less inclined to write imaginative stories.

Among the topics in task 7 the first three have been the most attractive. In grade 9 1/4 have written about what would happen if the pupils had more influence at school, whereas grades 2^2 and 3^3 above all discuss whether TV watching makes independent thinking more difficult. Perhaps these subjects are closer to the pupils' interests than the others, or maybe the latter are considered as more difficult to deal with in the short time available.

Mean scores and standard deviation

The calculations of the means of the compositions have been based on the marks that two raters - in the case of task 6 three raters - have given the composition. The mean of the raters' scoring have thus been calculated first of all. In the cases when one rater has given a mark for a composition whereas another has found it illegible or that the topic has been misunderstood, the task has been looked upon as worked out and the given mark has been used. The number of compositions that mean values and standard deviations in table 17 are based on, is therefore as regards several topics a little higher than what is stated in tables 13-15.

Mean values and standard deviations are given for overall impression and partial aspects except for legibility, that is handwriting, which will be accounted for later. In task 1E only overall impression marks were given.

The raters were instructed to read each composition at their normal reading speed and thereafter give a mark for the first impression that they received of the composition as a whole. This mark they were not to change later, when for example they had examined the content or the readability more in detail. The purpose of this appeal was to prevent the mark for overall impression from being a mean of the marks for the partial aspects.

The raters seem to have followed the instruction. Even if the mean differences are not great between the marks within the tasks, one can observe that the mean for an overall impression is almost regularly lower than for the partial aspects. The mean for content is higher than for overall impression in all the grades and tasks. There is also a higher value for disposition except for 1A in grade 9, for style except for tasks 1A and 1C in grade 9 and for readability except for task 5 in grades 6 and 9. It is only in these five cases that the mean value in a partial aspect is lower than the mean for an overall impression. In the others they are equal or in the majority of cases higher than the mean for an overall impression. The means are furthermore considerably much higher in some cases. In grade 9 the mean is .5 units higher for readability in task 1E and .4 units higher for content in task 1C.

It therefore seems to be a fact that the raters, when they study a composition from a certain aspect, irrespective of which, are more inclined to give it a higher mark than what they found the composition is worth at their first reading of it. It is difficult to explain this almost general difference. It is easier to realize the differences in means in task 1C. Both in grades 6 and 9 the pupils have managed better as regards content than style. Earlier on has been pointed out the difficulties the pupils have had in finding the right linguistic tone in this message to the headmaster. The same difference between content and style can be concluded from the means in task 1E.

It is also natural that it should be possible to find a difference between content and readability. The pupils may have something to tell but have formal deficiencies, which result in the mean being higher for content than for readability. This we find in the personal narrative which the compulsory school pupils are accustomed to and like to write. In the same way the pupils have managed better as regards the content of the message to the headmaster than the linguistic correctness. We cannot, however, find any such differences between aspects of content and form concerning the upper secondary school students.

Table 17. Means and standard deviations per task in grades 6, 9 och 2^2, 3^3

O = overall impression
C = content
D = disposition
S = style
R = readability

Task		Mark	Grade 6		Grade 9		Grades 22 ,33	
			M	S	M	S	M	S
1A	description of bicycle	O	2.8	1.0	2.7	.9		
		C	3.0	1.0	2.8	.9		
		D	2.9	.9	2.6	.9		
		S	2.9	.9	2.6	.9		
		R	3.1	1.0	2.7	.9		
1B	self-description	O	2.7	1.0	2.6	.9		
		C	2.7	1.0	2.7	.9		
		D	2.8	.9	2.6	.9		
		S	2.8	.9	2.7	.9		
		R	3.0	1.0	2.6	.9		
1C	message to the headmaster	O	2.5	.8	3.1	.9		
		C	2.8	1.0	3.5	1.0		
		D	2.7	.9	3.3	.9		
		S	2.5	.8	3.0	.9		
		R	2.6	.8	3.3	.9		
1D	message to the home	O	3.3	.9				
1E	advert answer	O			2.2	.9	3.0	.9
		C			2.4	1.0	3.3	1.0
		D			2.4	1.0	3.2	1.0
		S			2.2	.8	3.1	.9
		R			2.7	.9	3.4	.9
5	personal narrative	O	2.5	.8	3.2	1.0		
		C	2.7	.9	3.3	1.1		
		D	2.7	.8	3.3	1.0		
		S	2.6	.8	3.2	1.0		
		R	2.4	.9	3.1	1.0		
6	argumentative essay	O	2.1	.7	2.8	.9	3.3	1.0
		C	2.2	.8	2.9	.9	3.4	1.0
		D	2.1	.7	2.9	.9	3.4	1.0
		S	2.2	.7	2.8	.8	3.3	.9
		R	2.2	.7	2.8	.8	3.3	.9
7	reflective essay	O			2.2	.9	2.8	1.0
		C			2.2	.9	2.8	1.0
		D			2.3	.9	2.8	1.0
		S			2.3	.8	2.9	.9
		R			2.4	.7	2.9	.9
9	advisory letter	O			2.3	.9	3.0	.9
		C			2.4	1.0	3.1	1.0
		D			2.4	1.0	3.0	1.0
		S			2.6	.9	3.2	.9
		R			2.6	.9	3.2	.9

In tasks 1A and 1B the estimates in grade 6 and in grade 9 have been made according to separate scales. In the other topics, except for 1D, which was written only by grade 6, a common scale was used. Even for task 6, which was written by all, a common scale was used.

Throughout one can find an increase of mean with each higher grade. Thus for example in task 5 the mean is 2.5 for the pupils in grade 6 and 3.2 for the grade 9 pupils as regards the overall impression. The older the pupils are, the better compositions they write, which is to be expected. The increase of the mean between the grades is well over half a step on the scoring scale for each task.

As has earlier been pointed out the rater has not known anything that might influence the scoring about the student who wrote the composition. Neither sex nor age has been evident from the composition paper. The rater has only with the aid of the content of the composition been able to draw conclusions regarding for example the pupil's age. The higher means for the older students within a topic are due to the fact that the raters have considered their compositions to be better.

Some comparisons between the pupils' performances in different topics within a grade are possible only in a few cases, when the grade has been scored according to its own scale. One can thus find that grade 6 has received almost the same means for tasks 1A and 1B, i. e. 2.8 and 2.7, but has acquired better results for 1D with a mean of 3.3. But it is impossible to compare with the mean value for tasks 5 and 6, as the intermediate stage pupils in task 5 have been scored together with grade 9 and in task 6 they have constituted a joint group together with the others. The mean values for grade 9 can likewise only be compared regarding tasks 1A and 1B. The values are almost equal in the tasks, 2.7 and 2.6. In the case of the upper secondary school students no comparisons between topics can be made.

The means can be related to the international scale. In the section "Mean scores" the role of the calibration essays was discussed, and in diagram 1 there was an account of the raters' results in comparison with the international scoring. It turned out that for some topics the raters were above the international scale, in several of them below, but that in some they were completely in accord with international evaluation. In table 18 has been indicated the mean scores for all the pupils who wrote a task and the deviation from the international value that the raters have according to diagram 1.

Table 18. Mean scores and rater deviations from the international norm

Task		Mean value	Deviation
1A	description of bicycle (grade 6)	2.84	-.65
1A	description of bicycle (grade 9)	2.67	+.35
1B	self-description (grade 6)	2.71	-.45
1B	self-description (grade 9)	2.63	-.42
1C	message to the headmaster	2.80	-.24
1D	message to home	3.29	+.15
1E	advert answer	2.75	-.89
5	personal narrative	2.75	-.04
6	argumentative essay	2.65	0
7	reflective essay	2.58	0
9	advisory letter	2.55	-.37

The Swedish evaluation would need to be revised with the values given in the column called "Deviation". In the topics in which the deviation is negative, the marks should be raised so that the mean for, say, task 1C would be 3.04.

The possibilities of using the calibration essays as an instrument for correction have been discussed earlier. Their deficiencies were found to be so serious that they cannot be used to make international comparisons possible. It is therefore impracticable to make statements about Swedish pupils' achievements compared with international norms.

For the legibility mark three degrees of scoring were used, where in short 5 means that the handwriting is clear and easy to read, 3 that the reader finds it difficult to understand certain words and 1 that the reader has obvious difficulties in getting through the text. The mean values for the pupils' legibility marks in different grades are reproduced in table 19.

Tabell 19. Mean values and standard deviations of the legibility mark in grades 6, 9 och 2^2, 3^3

Grade	M	S
6	4.1	1.2
9	4.2	1.0
2^2, 3^3	4.4	1.0

A mean value above 4 implies that only a small number of pupils have got the mark of 1. They constitute well over 100 pupils per stage. Of the other the majority received the highest mark in practically all the tasks.

The percentage shares for the three degrees of marking differ from one task to the next. About 3/4 of the pupils in grades 9 and 2^2, 3^3 have received the highest mark in topics directed towards a specific reader, namely the message to the headmaster, the advertisment answer and the letter giving advice on how to write a composition. In the other subjects without a definite addressee the means in these grades are considerably lower. To what extent this difference is an accidental occurrance or the result of the pupils' ambition to write more neatly, when they address themselves to a particular person, is hard to have a definite opinion of. More or less unconsciously the pupils are perhaps influenced to be careful about their handwriting when they are thinking of a particular addressee.

The means get higher with each grade. The older the pupils are, the more flowing and legible their handwriting seems to be. This invalidates the view sometimes expressed that pupils' handwriting during school years gets worse and more and more characterized by carelessness and obscurity.

Correlations between pupils' achievements in different tasks

Some of the tasks are pragmatic as shown in chapter 2. The description of a bicycle, the self-description, the messages to the home and to the headmaster as well as the advertisment answer are such more practical writing tasks. Even the advisory letter can be considered to be practical.

The personal narrative is of a different character. The pupils are expected not only to tell about what they have experienced but also to describe how they took part in those events and express their reactions. The telling of the story usually runs in chronological order, irrespective of whether it refers to imaginative or experienced events, and the disposition therefore causes no difficulties at all.

To arrange the material can, however, be more complicated in the argumentative task. The pupils should not only express their views in a suitable order. They must also try to deliver their arguments intelligibly and convincingly, so that the reader is clearly aware of what the pupils wish to say. To find arguments, perhaps to find counter-arguments and to meet them in order to support one's own idea with purpose of convincing the reader is considerably more difficult than telling one's own experiences.

At least as demanding is the reflective task. The pupils are expected to offer their own thoughts and reflections in connection with the chosen headline. The quality of the composition is dependent on how rich it is in essential views and how these get developed and discussed.

In the written tasks different demands are thus put on the pupils. To what extent all the pupils succeed in solving a problem can be dependent on how much they have had an opportunity to practise writing it but also depends on their age and maturity. The question which will be elucidated in this section does however not refer to the achievements of the group of pupils as a whole, but instead the intention is to make a comparison between the achievements of the individual pupils dealing with different kinds of tasks. The pupils may well be expected to succeed better with certain tasks than with others.

To illustrate the concordance between the pupils' performances as regards different types of tasks the correlation coefficients have been calculated within the grades. In table 20 the correlations between the topics in grade 6 are stated. As tasks 1A-1D have been rotated as has earlier been described in chapter 2, and thus no pupil has written more than one of these tasks, no correlations can be calculated between these. In grade 9 1A, 1B, 1C and 1E have been rotated as well as 5, 6 and 7. In grades 2^2, 3^3 it is 1E and 9 as well as 6 and 7 that have been rotated. No agreement between the pupils' results in these topics can therefore be reported.

Tabell 20. Correlations between writing tasks in grade 6

		5	6	
1A	description of bicycle	.55	.50	
1B	self-description	.56	.50	
1C	message to the headmaster	.49	.47	
1D	message to home	.40	.50	
5	personal narrative	·	.54	
6	argumentative essay		·	M = .50

Almost all correlations are around .50, which is also the mean. In the next table the values for grade 9 are reported.

Table 21. Correlations between writing tasks in grade 9

		5	6	7	9	
1A	description of bicycle	.67	.61	.60	.61	
1B	self-description	.63	.70	.44	.61	
1C	message to the headmaster	.54	.53	.30	.45	
1E	advertisment answer	.60	59	.42	.55	
5	personal narrative	·	--	--	.59	
6	argumentative essay		·	--	.64	
7	reflective essay			·	.58	M = .56
9	advisory letter				·	

The mean value is somewhat higher for grade 9 and the distribution considerably wider than in grade 6. The lowest correlation is .30 and the highest .70. There is a moderately strong correlation between the tasks with correlation coefficients between .60 and .70.

Even for the upper secondary school students the correlation is around .50 as shown in table 22.

Table 22. Correlations between written tasks in grades 2^2, 3^3

		6	7	9	
1E	advert answer	.46	.50	--	
6	argumentative essay	·	--	.57	
7	reflective essay		·	.50	
9	advisory letter			·	M = .51

The majority of the correlation coefficients in the three grades are around .50 or just above that with certain diverging values in grade 9. A correlation of .50 implies that there is a good chance for the pupil to get the same or almost the same marks for the written tasks, that is a deviation of at the most one degree on the scoring scale.

Larsson (1984) in his study made the pupils write two compositions. One was designed as free writing with a picture used as a stimulus, and there was a purpose of directing towards relating and describing. The second composition was argumentative. Larsson received a correlation of .80 between the pupils' marks for the two tasks, which is considerably higher than in the IEA study. The higher value is, however, due to the fact among other things that Larsson used a multi-graded scale, which results in a higher agreement among the raters. In the section on rater agreements Larsson's results are discussed more in detail. Thus it seems to be a fact that the pupils achieve a general proficiency in writing. They manage to do different types of writing tasks on the whole equally well.

Such a conclusion must not, however, be taken as a pretext for the use of writing practice entirely aiming at a limited number of writing variants. Writing methodology should endeavour to train many different types of tasks. Even if the pupils may succeed in dealing with a task which is unusual for them it is very probable that they will manage even better, if they have opportunities for writing such exercises. Furthermore variations within their writing practice would certainly be felt as a welcome motivation.

Correlations between partial marks within topics

In the previous section on "Means scores and standard deviation" the overall impression mark and the partial marks within each task with the exception of the mark for legibility were reported in table 17. The differences between the means were generally some tenths of one degree on the scoring scale.

The correlations between the partial marks have been calculated for a number of tasks in each form. In table 23 there is a report on the agreement between the intermediate stage pupils' achievements as regards writing a message to the headmaster, writing a personal story and producing an argumentative essay.

Table 23. Correlations between partial marks for topics 1C, 5 and 6 in grade 6

O = overall impression
C = content
D = disposition
S = style
R = readability

		1C message to the meadmaster					5 personal narrative					6 argumentative essay				
Task		O	C	D	S	R	O	C	D	S	R	O	C	D	S	R
1C	O	·	.84	.82	.80	.74										
	C		·	.78	.74	.61										
	D			·	.77	.67										
	S				·	.66										
	R					·										
5	O						·	.89	.86	.89	.82					
	C							·	.85	.81	.69					
	D								·	.82	.73					
	S									·	.83					
	R										·					
6	O											·	.87	.86	.83	.77
	C												·	.82	71	.62
	D													·	.77	.70
	S														·	.85
	R															·

The correlations between the overall impression and the partial marks of content, disposition and style are throughout high and around .85. The correlations between the formal skills, that is readability, and the totality are not equally strong. They are weakest between the aspects of content and disposition on the one side and the grammatic and orthographic skills on the other.

This outcome can be compared with the results that Page and Paulus (1968) received from profiled evaluation. The correlations between content oriented aspects were high but only a little more than .65 between content aspects and linguistic correctness.

As is shown in table 24 the same patterns returns in grade 9. The correlations are moreover somewhat higher. Earlier it has been indicated that the pupils in general found it hard to acquire a suitable linguistic tone in the letter to the headmaster, even if the content of the message did not cause any trouble. This condition can also be seen in the weaker correlations both in grade 6 and grade 9.

Table 24. Correlations between partial marks for topics 1C, 5 and 6 in grade 9

O = overall impression
C = content
D = disposition
S = style
R = readability

		1C message to the meadmaster					5 personal narrative					6 argumentative essay				
Task		O	C	D	S	R	O	C	D	S	R	O	C	D	S	R
1C	O	·	.83	.86	.87	.80										
	C		·	.82	.75	.65										
	D			·	.82	.75										
	S				·	.73										
	R					·										
5	O						·	.92	.91	.92	.87					
	C							·	.90	.85	.78					
	D								·	.87	.83					
	S									·	.88					
	R										·					
6	O											·	.93	.92	.87	.85
	C												·	.89	.82	.77
	D													·	.86	.81
	S														·	.88
	R															·

Tasks 1C and 5 were not written by the pupils in grades 2^2, 3^3. Therefore the correlations between the partial marks of these pupils for tasks 6, 7 and 9 are given. The correlations are as shown in table 25 even higher than in the two lower grades. Correlation between readability and other partial aspects are also here lower than between content, disposition and style.

Table 25. Correlations between partial marks for topics 6, 7 and 9 in grades 2^2, 3^3

O = overall impression
C = content
D = disposition
S = style
R = readability

		6 argumentative essay					7 reflective essay					9 advisory letter				
Task		O	C	D	S	R	O	C	D	S	R	O	C	D	S	R
	O	•	.95	.93	.90	.87										
	C		•	.91	.86	.80										
6	D			•	.88	.84										
	S				•	.91										
	R					•										
	O						•	.92	.91	.85	.83					
	C							•	.87	.80	.74					
7	D								•	.81	.77					
	S									•	.87					
	R										•					
	O											•	.90	.88	.86	.79
	C												•	.83	.79	.71
9	D													•	.80	.74
	S														•	.82
	R															•

The correlations are in all grades so high that in all calculations that refer to the pupils' achievements it is possible to use the overall impression mark. Only the readability mark correlates somewhat lower with the overall mark.

It is, however, possible to find that the raters have varied in their scoring of content and readability. In grade 6 the correlation between these partial marks is around .65 and in the other around .75.

Attitudes towards writing tasks

In connection with the writing task the pupils were asked to answer three questions of attitude. These are reproduced in chapter 2.

Means and standard deviations of the answers have been calculated and are shown in table 26. For question A1 the mean can vary from 1.0 (not common at all) to 3.0 (very common) and for questions A2 and A3 1.0 (difficult and boring respectively) to 5.0 (easy and amusing respectively).

Table 26. Means and standard deviation of questions of attitude nos. 1-3 for written tasks in grades 6, 9 and 2^2 and 3

Task	Grade 6						Grade 9						Grades 2^2, 3^3					
	A1(us)		A2(easy)		A3(fun)		A1(us)		A2(easy)		A3(fun)		A1(us)		A2(easy)		A3(fun)	
	M	s	M	s	M	s	M	s	M	s	M	s	M	s	M	s	M	s
1A bike	1.3	.5	3.7	1.0	3.7	.9	1.2	.4	3.5	1.1	3.0	1.1						
1B self-descr.	1.3	.6	3.6	1.1	3.8	1.0	1.3	.5	3.5	1.1	3.2	1.1						
1C mess. headm.	1.4	.6	3.5	1.1	3.6	1.0	1.3	.5	3.5	1.0	3.1	1.0						
1D mess. home	1.7	.6	4.3	.9	3.8	.9												
1E advert							1.2	.5	2.9	1.1	3.1	1.1	1.2	.5	3.1	1.1	3.0	1.0
5 pers. narr.	1.7	.6	3.4	1.2	3.7	1.2	1.6	.6	3.0	1.2	3.2	1.3						
6 argum.	1.1	.4	2.7	1.3	3.1	1.3	1.1	.3	2.3	1.2	2.7	1.3	1.4	.6	2.5	1.2	2.7	1.2
7 refl.							1.2	.4	2.3	1.2	2.6	1.3	1.6	.6	2.6	1.1	2.5	1.2
9 advice							1.1	.3	3.0	1.3	2.7	1.2	1.1	.3	3.2	1.1	3.0	1.1

The majority of the pupils are of the opinion that the written tasks are not at all usual in their teaching. For most of the tasks the means are only a few tenths above 1. The intermediate stage pupils have the highest value for the personal narrative and the message to the home. Probably their number of answers referring to the latter rather shows how often they write such messages at home than its presence during teaching hours. The instruction does not indicate either that the questions concerns writing only at school. In grade 9 the personal story has the highest value and in the upper secondary school the argumentative and reflective tasks, which seems most likely. The low means for the argumentative task in grades 6 and 9 as well as the advisory letter in grades 9 and 2^2, 3^3 are also what could be expected.

The pupils' answers can be compared with their estimates in the pupil questionnaire of the question how often they write different kinds of tasks (Löfqvist, 1988b). To make up their own stories and write about personal experiences are common writing tasks in the compulsory school. On the other hand, the pupils are seldom allowed to write about their reflections on what they have perhaps seen, heard or read, and not to describe problems that should be solved and then also suggest solutions and argue in favour of them.

These kinds of topics are however more common in the upper secondary school.

In their questionnaire the teachers have been asked to estimate how often they use different writing tasks. They say that they seldom or never offer the task of writing an advisory letter (Löfqvist, 1988b). It is therefore not surprising that the senior stage and upper secondary school pupils state that topics like advice on composition writing are unusual, and that they therefore have low mean values.

The differences between the means are otherwise small for question A1, that is concerning how common the pupils consider the tasks to be. But they point in the same direction as the teachers' and the pupils' estimates of how often different topics are written about at school. Higher and lower means in the different grades correspond to high and low frequencies in the teacher and pupil estimates.

All pupils think that the argumentative task is the most difficult one. It has got the lowest mean for question A2 in all age groups. The intermediate stage pupils mean that the other tasks are quite easy and then especially the message to the home. The argumentative task is, however, one step on the scale lower than the others.

The senior stage and upper secondary school pupils do not consider the other tasks either difficult or easy. Their means are around 3.0 except for the shorter task, which grade 9 finds rather easy. On the other hand the older pupils have lower means than grade 6 for the argumentative task. It may seem peculiar that the older pupils, of whom in any case the upper secondary school students practise argumentative essays, consider it more difficult than the intermediate stage pupils to whom such a task is more unfamiliar. The explanation is probably not that the older pupils would understand better the demands such a task puts on the writer, but rather that this difference reflects the general attitude towards writing among the pupils. A greater proportion of grade 6 than of the others like to write. At the intermediate stage 2/3 of the pupils say that it is fun or great fun to write compared with around half the pupils at the other stages (Löfqvist, 1988b).

This attitude towards writing among the pupils in grade 6 we also find as regards question 3. Their attitude to all the written tasks is positive. The mean is above 3.0 even for the argumentative task, which they consider to be more unusual and more difficult than the other tasks. The intermediate stage pupils are also on average more positive to the written tasks than the older pupils. The mean value of means for question 3 is 3.6 in grade 6, 3.0 in grade 9 and 2.8 in grades 2^2 and 3^3.

Correlations between the attitudinal questions

Throughout it seems as if a higher mean for the second question of attitude corresponds to a higher value for question 3, that is if the majority of the pupils have considered the task rather easy or easy, then the majority have also considered it rather amusing or amusing. To get a measure of how strong the agreement is between the pupils' answers to the three questions, the correlations have been calculated. They are given in table 27.

Table 27. Correlations between attitudinal questions 1-3 for writing tasks in grades 6, 9 and 2^2, 3^3
A1 (common) A2 (easy) A3 (funny)

		Task									
		1A	1B	1C	1D	1E	5	6	7	9	**M**
Grade 6	A1 - A2	.18	.12	.13	.12	-	.21	.20	-	-	**.16**
	A1 - A3	.11	.16	.13	.09	-	.21	.18	-	-	**.15**
	A2 - A3	.22	.39	.30	.12	-	.50	.54	-	-	**.35**
Grade 9	A1 - A2	.04	.18	.12	-	.20	.31	.22	.20	.10	**.17**
	A1 - A3	.12	.11	.12	-	.12	.27	.17	.15	.09	**.14**
	A2 - A3	.34	.36	.25	-	.27	.52	.50	.51	.49	**.41**
Grades 2^2, 3^3	A1 - A2	-	-	-	-	.19	-	.27	.28	.03	**.19**
	A1 - A3	-	-	-	-	.08	-	.17	.12	.07	**.11**
	A2 - A3	-	-	-	-	.17	-	.40	.41	.33	**.32**

The correlations between A1 and the other attitudinal questions are positive but weak throughout. That a writing task is unusual thus does not imply, that it is looked on as either difficult or easy, nor as dull or funny. There is, however, a certain correlation between frequency and degree of difficulty in grade 9 for task 5 and in grades 2^2, 3^3 for tasks 6 and 7 indicating that frequent task are regarded as easy.

There is a better agreement between the students' answers to questions 2 and 3. This is evident especially as regards the longer compositions namely that a task which is considered difficult is also looked on as dull, whereas the tasks that the pupils find easier are also estimated as more fun to write. The strongest correlation is to be found in the longer compositions 5, 6, 7 and 9.

Correlations between scores and attitudes

It is of great interest to examine, if there is a correlation between the pupils' performance and attitude towards the task. One would for example be inclined to think that the pupils would succeed better, that is receive better marks, for such types of compositions as they are accustomed to writing than for written tasks that they have not practised before. What the correlations are between scores and attitudinal questions is shown in table 28. When these calculations were made the overall mark was used.

Table 28. Correlations between scores (B) and attitudinal questions for written tasks in grades 6, 9 och 2^2, 3^3
A1 (common) A2 (easy) A3 (funny)

		Task									
		1A	1B	1C	1D	1E	5	6	7	9	M
Grade 6	B - A1	.07	.09	.12	.04	-	.13	.06	-	-	**.09**
	B - A2	.10	.25	.17	.14	-	.12	.12	-	-	**.15**
	B - A3	.16	.27	.11	.18	-	.21	.23	-	-	**.19**
Grade 9	B - A1	.00	.16	.10	-	.06	.23	.06	.04	-.01	**.08**
	B - A2	.12	.31	.12	-	.20	.27	.17	.26	.29	**.22**
	B - A3	.26	.43	.23	-	.27	.42	.35	.44	.44	**.36**
Grades 2^2, 3^3	B - A1	-	-	-	-	.11	-	.26	.23	-.07	**.17**
	B - A2	-	-	-	-	.17	-	.18	.24	.28	**.22**
	B - A3	-	-	-	-	.20	-	.32	.27	.28	**.27**

There is no correlation between scores and the answers to question A1. The correlations are low throughout. It is thus not the case that the pupils' performances vary due to how common the written tasks are at school. Nor does it seem to be of any importance to the performance if the pupils consider the task to be difficult or easy. The correlations are positive but weak. Somewhat higher is the correlation with question A3, but the quality of the mark is still rather independent of whether the pupil considered it amusing or boring to write. Even in those tasks where the correlations are highest predictions of the pupils' performances considering how much they appreciated the tasks are uncertain.

Mean scores of the judges

The scorers' means have been discussed in different contexts earlier on. In the section "Mean scores" the values of the teachers for two sets of compositions which had also received an international evaluation, namely the benchmark set and the calibration set, were accounted for. A comparison was made between the national and the international evaluation.

In the same section were shown on the one hand the results of a comparison between the teacher groups' means for the same tasks, on the other the results of a repeated scoring of the same compositions. No differences were to be found between the teacher groups. The teachers were also constant in their scoring.

All these calculations have been made on smaller numbers of compositions. The calibration essays were thus on average 7 to 8 in each task, in the benchmark set were 10 to 12 compositions and the repeated ones were 20 for each topic. All raters within one topic have marked these compositions and it has thus been possible to compare them with each other.

In the main investigation each composition has been evaluated by two teachers except for the argumentative task, which had three raters. As the raters are coupled together according to the model described in paragraph 5.3.1, comparisons can be made between more than two raters within a task. Such comparisons will be made in this section.

The distribution of the compositions according to paragraph 5.3.1 to different bundles was made at random. The total number of compositions per topic which each rater had to score, was according to table 32 between 110 and 340. As for each rater the share of compositions within a topic was large, there is reason to believe that those shares did not differ from each other as regards quality. Each marker has thus had to work on equally good bundles of compositions.

In following tables the differences between the raters' means for the overall impression in each task will be illustrated. The highest mean for a task has been indicated with an H and the lowest with an L. Furthermore the mean of the scorers' means, the standard deviations and the number of evaluations (n) are shown. The raters are indicated by code numbers, in which the hundred figure indicates teachers at the intermediate stage (1), the senior stage (2) or the upper secondary stage (3). Table 29 comprises the tasks that have been evaluated by means of separate scales.

Table 29. Means per rater for tasks 1A (description of bicycle), 1 B (self-description) and 1D (message to the home) for grades 6 and 9

Rater	Grade 6			Grade 9	
	1A	1B	1D	1A	1B
103	2.9	2.7	3.3		
106	2.8	2.8	3.3		
107	3.1 H	2.9 H	3.3		
112	2.8	2.7	3.3		
113	2.6 L	2.5 L	3.3		
211				2.5 L	2.7 H
212				2.6	2.6
213				2.6	2.7 H
216				2.9 H	2.5 L
217				2.8	2.7 H
M	2.8	2.7	3.3	2.7	2.6
n	579	539	590	821	799
s	1.0	1.1	1.0	1.0	1.0

The most significant agreement can be found between the markers in task 1D in grade 6. They all have a mean of 3.3. In the other tasks in grade 6 there is a difference of about half a step on the scoring scale between the raters. Number 107 is the most lenient and number 113 the strictest. The differences between the senior stage teachers is a little smaller namely .4 steps in task 1A and .2 in 1B.

By means of the n-figures one can estimaste how many compositions that each marker has taken care of. The figures were divided by the number of markers, which results in about 100-125 compositions per topic in grade 6 and about 160 in grade 9.

In table 30 the tasks that were written by pupils in two grades and which were estimated according to a common scale have been put together. The number of raters has been ten for tasks 1C, 5 and 7 and twenty for 1E and 9. Task 6 is shown in table 31.

The differences between the raters' means are greater than what we observed as regards the tasks that were accounted for in table 29. It is well over half a step for tasks 1C and 5 and around one step in 1E, 7 and 9. Thus we find that the differences between the raters get larger in those tasks, where the number of raters is the largest, which is natural, as the possibilities are then greater of getting diverging values. Also in task 6 with 27 raters the difference is one step which is shown in table 31.

Table 30. Means per rater for tasks 1C (message to the headmaster), 1E (advert answer), 5 (personal narrative), 7 (reflective essay) and 9 (advisory letter)

Rater	1C (grades 6,9)	1E (grades 9, 22, 33)	5 (grades 6, 9)	7 (grades 9, 22, 33)	9 (grades 9, 22, 33)
101	2.8		2.9		
102	2.8		2.7		
103	2.6		2.5 L		
110	3.1 H		2.7		
111	2.9		3.1 H		
201	2.6	2.7	2.7		2.3
202	2.9	2.5	2.7		2.1
203	2.9	2.9	2.8		2.8
204	2.9	2.5	3.0		2.3
205	2.5 L		2.6		
206		2.9		2.6	2.7
207				2.8	
208		3.1		3.0	2.8
209				2.4	
210				2.1 L	
214		2.6			2.5
215		3.2			2.6
218		2.7			2.5
219		2.4			2.7
220		3.4 H			3.0 H
302		2.8		3.1 H	2.7
303		2.5		2.4	2.5
304				2.5	
305				2.5	
306				2.4	
310		3.4			2.8
311		2.8			2.8
312		2.8			2.0 L
313		2.6			2.4
314		2.5			2.2
315		2.3 L			2.6
316		2.8			2.7
M	2.8	2.7	2.8	2.6	2.6
n	1401	2219	3386	2042	4537
s	1.0	1.1	1.0	1.1	1.0

Table 31. Means per rater for task 6 (argumentative essay)

Rater	6 (grades 6, 9, 2^2, 3^3)	Rater	6 (grades 6, 9, 2^2, 3^3)
101	2.7	208	2.7
102	2.7	209	2.6
103	2.6	301	2.3 L
104	2.6	302	2.9
105	2.8	303	2.5
106	3.3 H	304	2.4
107	3.0	305	2.5
108	2.8	306	2.4
109	2.7	307	2.5
201	2.4	308	2.8
202	2.9	309	2.6
203	2.8		
204	2.5	M	2.7
205	2.5	n	6831
206	2.6		
207	2.4	s	1.1

The growing difference with an increasing number of scorers is illustrated in table 32. The number of compositions per rater is varied but large. It is therefore not probable that, which has been discussed earlier, a diverging mean can be caused by the fact that one rater has received compositions that were thought to be exclusively above or below average. The difference in steps on the scoring scale within a task is dependent on the number of raters.

Table 32. Difference in average mark values between raters

Task	Diff on scoring scale	Number of raters	Number of compositions per rater
1A (grade 6)	.5	5	ca 115
1B (grade 6)	.4	5	" 110
1D (grade 6)	.0	5	" 120
1A (grade 9)	.4	5	" 160
1B (grade 9)	.2	5	" 160
1C	.6	10	" 140
5	.6	10	" 340
7	1.0	10	" 200
9	1.0	20	" 225
1E	1.1	20	" 110
6	1.0	27	" 250

The differences between the raters that we have found in the IEA study can be compared with the results that Eneskär got when she compared the raters at the standardization of the tests in Swedish for the compulsory school (Eneskär, 1988). Five scorers of 150 composition chosen at random nearly all are around 3.0. The study refers to evaluations made during four years. The greatest difference is .3 steps on the scoring scale. This agreement between the raters is, however, dependent on the fact that they had been instructed to assort the compositions according to the percentage indications of the five-graded scoring scale, which ought to have led to mean of 3.0 for all, if the instruction had been followed. In the IEA study, as the guide for the scorers shows, the instruction has meant that the teachers should not "adjust or in any other way adapt the proportions of the number of compositions with different marks, for example by using a distribution according to the normal curve". The compositions that each teacher had to evaluate were only part of a large number.

In the FRIS project (Martinsson, 1978, p 85) it says that "the compositions should be evaluated like ordinary school compositions". Probably this instruction has also meant that the raters were to aim at a distribution of the marks according to the percentage indications of the curriculum, as the mean for practically all the scorers are around 3.0 with a standard deviation of 1.0. Larsson (1984, p 143) in his instruction for the scorers has stated that they ought to make use of their teaching experience "but at the same time glance at the normal distribution curve. Within each stage there should be ones as well as fives". The raters' means vary considerably in spite of this instruction. Although the number or raters in different grades is not higher than 6 or 7, the variation is around 1 step on the scoring scale between the highest and the lowest mean value.

The differences in means in the IEA study do thus not seem remarkable. Rather do they seem to illustrate the variation in demands that characterizes teachers' scoring and which obviously it is not possible to eliminate in spite of special training with access to a benchmark set.

In the introduction to this praragraph on the scorers' means we told about the earlier calculations of the means for different sets of compositions. These calculations were intended to check, whether the teachers had in their scoring adjusted themselves to the international scale. The calculations, which had been carried out on the extensive material of the main investigation were meant to make it easier for us to estimate the reliability of the teachers' scoring with more certainty than during the check-up. We have found here that the difference between the means may amount to about one step on the scoring scale, when around 20 teachers are used as raters.

A difference between two raters can be illustrated by means of the following table. In one task 60 scored compositions have been chosen at random. The marks are given for each composition. The mean for the first teacher is 3.1 and for the other 2.4.

No.	Mark	No.	Mark	No.	Mark	No.	Mark	No.	Mark	No.	Mark
1	2-1 +	11	2-1 +	21	3-1 +	31	1-1	41	2-1 +	51	2-1 +
2	3-2 +	12	5-5	22	4-3 +	32	3-1 +	42	3-4 -	52	5-3 +
3	3-4 -	13	2-4 -	23	3-4 -	33	3-2 +	43	3-1 +	53	3-2 +
4	3-1 +	14	3-3	24	2-2	34	1-1	44	3-3	54	5-5
5	2-2	15	4-3 +	25	4-2 +	35	2-2	45	3-2 +	55	3-2 +
6	5-2 +	16	1-1	26	3-1 +	36	5-4 +	46	3-3	56	4-3 +
7	1-2 -	17	5-4 +	27	3-1 +	37	1-1	47	3-2+	57	3-2 +
8	3-2 +	18	5-4 +	28	5-3 +	38	2-1 +	48	1-1	58	3-3
9	5-4 +	19	5-3 +	29	3-1 +	39	3-4 -	49	3-3	59	1-1
10	5-4 +	20	3-1 +	30	3-3	40	3-4 -	50	5-3 +	60	2-2

Plus and minus signs indicate that the first teacher has the highest and the lowest marks for the composition respectively. In the case of 35 compositions the first teacher has a more generous estimate but in 7 cases he is more restrained. The two teachers give the same marks for 18 compositions. The difference is greatest as regards essay 6. The first rater is here 3 steps higher but in the case of number 13 he is 2 steps lower.

A mean difference consisting of .7 steps, as in this case, may thus imply considerable differences of scoring concerning individual compositions, as the differences between the raters spread in different directions. Usually the difference is one step on the scale, but for 1/4 of these compositions it is two or several steps.

We are coming back to the question of agreement between evaluations in the next paragraph.

Correlation between the judges

The correlation between the raters has earlier been calculated during their scoring, of among other things the calibration and repetition compositions. The results of these calculations have been accounted for in "Correlation between the scorers". Table 10 shows that the mean of their correlations for these compositions is .75.

In the main investigation the raters have been coupled together in pairs except for the argumentative task, in which each composition has been scored by three raters. Two raters may in certain cases have scored several different tasks together. That is for example the case in grade 6 as regards tasks 1A the bicycle, 1B the self-description and 1D the message to the home, which have all been marked by five teachers of the intermediate stage. The personal story was written both by grade 6 and grade 9 and was evaluated by means of a common scale. Teachers from both the intermediate and senior stages were used as judges, and each composition was evaluated by one teacher from each stage. In the combinations of judges that were created for the tasks, it has been an ambition to design as many different combinations of judges as possible so that they should in common have evaluated as many tasks as possible of different types.

In the following table there is an account of the number of correlations between raters, divided according to grades and tasks, that were estimated accoring to a separate scale (1) or a common scale for two (2) or three (3) grades. The correlation coefficients are noted in intervals with a class-interval of .05.

Table 33. Number of correlations between raters per grade with the correlation coefficients distributed into intervals

1 = separate scale for each grade
2 = common scale for two grades
3 = common scale for three grades

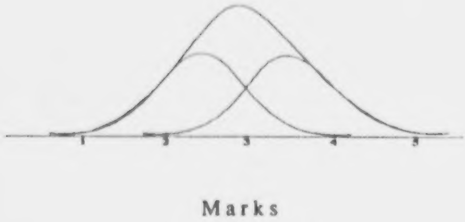
	Grade 6			Grade 9			Grades 2^2, 3^3	
	1	2	3	1	2	3	2	3
0 - .05					1			

.36 - .40			1		1		1	
.41 - .45			1		1			
.46 - .50		1			4	1	2	2
.51 - .55	1	3	1		4	2	3	
.56 - .60	1	6	7		5	3	10	1
.61 - .65	3	7	7		7	6	8	6
.66 - .70		3	4	1	17	7	11	6
.71 - .75	4	6	4	5	16	1	12	2
.76 - .80	1	2	2	3	15	5	5	8
.81 - .85		1		1	13	1	5	2
.86 - .90		1			3	1	1	
M	.67	.66	.63	.75	.67	.67	.67	.69

The number of compositions that two teachers have evaluated together within one grade may vary from exceptionally well over ten up to several hundred. Usually the correlations are based on some 50 compositions.

The correlations are lower than those that were accounted for in the section "Correlation between the scorers". The means according to table 33 is between .63 and .75. These values correspond very well with what one usually gets in similar investigations. The explanation given earlier for the higher correlations during the check-up of the raters therefore seems plausible.

In the discussion of the means in connection with table 17 we found that they were highest in the higher grade in tasks which were evaluated according to a common scale. The judges have thus given higher marks to the older pupils, whereas the younger ones have as a rule received the lower marks. The scoring scale for one grade in a common task has been compressed in a way that can in principle be illustrated with the following figure. The smaller curves indicate the distribution of marks of two groups of pupils within the overall distribution.



It often turns out that a limited part of the scoring scale is used for each grade and an even more restricted part if three grades are estimated according to the same scale. The whole scale is used for one grade only if it is separate for the grade.

The correlations between judges are lower, if only part of the scale is used and higher, if all the steps of the scale are made use of. In table 33 one can find such a tendecy towards an increase of the correlation coefficients except for grades 2^2, 3^3. There, on the other hand, the differences between the mean correlations are unimportant.

The marks for the 60 compositions which in the previous section illustrated a difference in mean of .7 steps on the scoring scale between two judges can be registered in a correlation table.

		Second teacher					
		1	2	3	4	5	
	5		1	4	5	2	12
	4		1	3			4
First teacher	3	8	8	6	5		27
	2	5	4		1		10
	1	6	1				7
		19	15	13	11	2	

The correlation coefficient is .65, that is about the same agreement that we have usually found between the raters according to table 33.

In Eneskär's (1988) report on the standardized achievement tests the average correlations for the five raters turned out to be between .63 and .70. The last time an investigation was carried out she tested a standardization using a 15-graded scale instead of the 5-graded one, as the raters hade made use of the possibility of giving plus and minus signs in addition to pure numerical marks. The mean correlation then rose to .76.

In Larsson's (1984) investigation the average correlations in different grades were between .65 and .73, when the 5-graded scale was used. He has, however, got considerably higher values in a sample for intensive treatment. From different grades he has chosen for boys and girls the six best, of medium quality and of the lowest standard of the compositions whose evaluations have been based on marks given by several judges. By doing that Larsson has designed a sample with a large number of extreme compositions as was the case with the compositions in the benchmark set and the calibration essays in the IEA study, which led to unusually high correlations between the raters. In Larsson's intensive sample there are also compositions from grade 7 up to grade 2^3 in the upper secondary school. The compositions have then been scored according to the 9-grade stanine scale, whose percental distribution were to be followed. As this procedure implies partly that a large number of extreme compositions have been used, partly that the raters have been forced to a solution with a certain number o compositions per step on the scale, the result must be a closer agreement between the raters. The result, however, only ostensibly implies a better agreement between the judges.

In the FRIS project (Martinsson, 1978) partly teachers, partly creativity researchers have had to work as scorers. The average correlations for the teachers were .71 in grade 4 and .72 in grade 6. The researchers were to estimate the compositions with special regard to their creativity, and they were allowed to design their own evaluation norms. The average correlation for them was .52 and .69. In "Swedish measures" (Hersvall et al, 1974), which worked on a debating article in the written central test, in addition to three teachers also three journalists were asked to evaluate the compositions. The average correlation for the teachers was .56 and for the journalists only .18.

Hultman and Westman (1977) made professionals carry out an evaluation of what they called "every-day Swedish", namely brochures, newspaper text and texts in school books. Each text was scored according to a 3-graded scale in the aspects of common interest, information density, disposition-presentation and linguistic clarity-comprehensibility. The purpose was that they should estimate each text in relation to their experience of that type of text in general. The professionals were as regards the brochure texts three writers at the National Board for Civic Information, concerning the newspaper texts a journalist, a journalist teacher and a teacher of Swedish and finally for the school book texts two educational editors and one teacher of Swedish. The correlations between the judges was however weak. For the brochure texts the correlations were .09, .27 and .47, for the newspaper texts .31, .41 and .49 and lastly for the schoolbook texts -.28, -.11 and .59.

It would seem as if correlations around .70 would be what can be expected as a measure of agreement between teachers. Björnsson (1960) in his calculations of scoring agreement had the value of .70, and he considered this normal for three judges. It also seems to be true that thanks to their education and experience teachers have better qualifications than others outside school to reach unanimous evaluations.

We must in addition to that state that the differences between the teachers' scoring of the pupils' compositions seem to be equally great, irrespective of whether the teachers make a traditional estimate or are specially trained for an analysis and evaluation of partial aspects in the compositions. Can we accept this subjectivity in their evaluation or can we find other ways for a more objective evaluation?

5.3.4 Language analysis and written achievement

Theoretic and empiric background

In the IEA project great pains have been taken to create an evaluation of the pupils' achievements that would make it possible to compare the results of students in different countries. In each country the same evaluation methods have been used with internationally constructed sets of compositions as a means to apply a reliable estimate.

The Swedish teachers who scored the compositions, have had special training in the use of the international scale, just like the judges in other countries, and they have had a detailed guide at their disposal. The results of the evaluations in the Swedish main investigation, however, do not indicate that the agreement between the teachers is better than the one usually attained between teachers that have had no special training and who have marked compositions in a conventional way.

At examinations comparable to the IEA study of the students' writing profiency the subjective evaluation procedures are both laborious and expensive. It is therefore interesting to analyse if in studies of this kind it is possible to substitute the human evaluation with other types of measuring of the pupils' achievements.

Some investigations that have used mechanical language measures to calculate linguistic proficiency are reported in "Swedish measures" (Hersvall et al, 1974). The before-mentioned researchers Page and Paulus (1968) have calculated, with the use of computer analysis of compositions, the correlation between marks on compositions, given by four independent judges, and different language measures. The strongest correlations was found with length of words (.51), number of frequent words (-.48), number of commas (.34) and number of words (.32). If all the language measures put together were included, the multiple correlation was .71, which was considerably higher than the inter scorer correlation (.50). Hunt (1964) made no comparisons between rater and language measure but compared the productivity of pupils of the same intellectual level that could be found in different grades. He found that the sentences got longer and that the language was characterized by greater density with increasing age. Older students used prepositional phrases to a greater extent and more complicated constructions, which resulted in a more mature language.

In the STIG-project (Pettersson, 1982) new writing exercises for the upper secondary school students aimed at making them write a more compact language, that is making them say the same thing with the use of fewer words. Transformational practice encouraged the students to use longer words and thereby to get closer to adult language.

In the two projects FRIS and "Swedish measures" the method of analysing the language with the help of computers has been further developed. FRIS (Martinsson, 1978) has among other things studied the development and evaluation of the language of free writing. A survey has been made of the students written language by means of a linguistic analysing system which was worked out within the projects "Speech syntax" and "Writing syntax"

(Teleman, 1974). The system is called MAMBA (Manual of sentence structure analysis). The analysis comprises sentences in the students' compositions which are analysed syntactically and lexically. For each word in a sentence it is decided which type of sentence it is part of, what part of speech it belongs to and what part of the sentence it is. The analysis resulted in 87 different measures of language, e.g. number of sentences, clauses, number of words belonging to parts of speech of different kinds, long and short words etc. In the following table the language measures are shown with the highest correlation with the teacher evaluations.

Table 34. Correlations between language measures and judgements (from Lindell, 1980, p 25)

Grade 4		Grade 6	
Different words	.75	Different words	.79
Punctuation marks	.74	Words used once	.77
Adverbials in S1	.74	Lines	.75
Lines	.73	Words	.75
Words H-G 1-200	.73	Words Allén 1-200	.74
Adverbials p.s.	.73	Words H-G 1-200	.73
Words Allén 1-200	.73	Adverbials p.s.	.72
Words	.73	Adverbials in S1	.72
Different verbs	.72	Punctuation marks	.72
Different adverbs	.72	Punctuation marks separating sentences	.71

"Words Allén 1-200" refers to the 200 most common words in Allén's frequency word list from 1972. It is based on newspaper texts whereas "H-G" (Hassler-Göransson) refers to an older frequency list from 1956, which had different texts as its basis. "Adverbials in S1" means that the adverbial consists of one word, "adverbials p.s." that the adverbial constitutes a "primary" part of sentence, irrespective of whether it consists of one or several words.

In the two grades the correlations between the teacher evaluations and the language measures are high. The latter are also nearly the same in both grades. The students' productivity, e.g. in the form of different words and number of words in the compositions, is a dominating factor and probably plays an essential part at the evaluation of their achievements.

During the comparison of the students' language in the two grades it turned out that grade 6 had more of everything. This is naturally a consequence of the fact that the older students have written longer compositions.

"Swedish measures" (Hersvall et al, 1974) used compositions that were written as standardized achievement tests in grade 3[3] in the upper secondary school. These have partly been evaluated in the conventional way by among others teachers, partly been analysed by means of the same linguistic system as in the FRIS project. The material has been factor analysed. The first two factors were dependent on the length of the compositions. The number of words had high loadings in these factors.

The language measures directed towards productivity turned out to have the highest correlation with the teacher evaluation. It was considered sufficient to use the frequency measures of different words, length of words and words in order to reach the best correlations with the teacher evaluation, and the values were understood to indicate "that it is safer to make a purely mechanical estimate by means of computer measures than to use individual judges, but it is somewhat better (though not much!) to make use of scoring groups of three persons".

The question of the correlation between marks for compositions and number of words in the compositions is discussed by Hultman and Westman (1977). They had divided the words of the compositions into two groups according to frequence. Up to and including mark 4 there was an increase in the compositions both of the number of "common" words and "uncommon" words. After that there was above all an increase of the more uncommon and pregnant words. The compositions of the fives were even shorter than those written by the fours. In the discussion of word count used as a scoring method Hultman and Westman mean that the number of words does not "contribute as much to the prospect of getting a five as to the risk of receiving a one" (p 54). They maintain that quantity is more important when one has to distinguish the poorer compositions than the better ones.

In Larsson's investigation (1984) the students got two tests in word knowledge. The correlations between the results on these and the students' written achievements were just above .70, and even stronger was the correlations between marks on the one side and the number of words, different words and long words on the other.

The correlation between the length of the composition and its mark was shown by Pettersson (1980) in an investigation of the 1977 standardized achievement test in Swedish in grade 3 of the upper secondary school, based on a sample of around 400 compositions on all the topics within the test. The number of words increases continuously from about 500 for the mark of 1 up to around 900 for the mark of 5.

The results of the investigations reported here show that productivity is closely related to the evaluation of written achievements. The calculcations have, however, been based on computer analysed compositions. To punch the compositions in an investigation is a both laborious and expensive procedure.

It would be much more convenient if an estimate of language measures could be carried out in a simpler way. Would it for example be possible in the IEA study to substitute the teachers' scoring of the compositions with a more simplified procedure, e.g. a count of the number of words and sentences in the compositions? Would it be possible with the use of such language measures to attain the same reliability in the estimate of the students' achievements as when teachers are asked to evaluate?

Word counting as a measure of achievement

As a preliminary study of the possibilities of using language measures a number of calculations were made. In task 6, the argumentative essay, the compositions in the benchmark set and the calibration set are in all 22. They have all been internationally scored. The correlations have been calculated between the international marks on the one hand and number of words, number of long words and number of punctuation marks on the other. These are shown in table 35. The mark for an overall impression has been used. As long words are indicated those that have seven or more letters. All punctuation marks are included apart from commas.

Table 35. Correlations between international mark and language measure for task 6

Language measure	Correlation with international mark
Words	.80
Long words	.73
Punctual marks	.62

The correlations between language measures and marks are high and then above all as regards the number of words. They are also higher than the correlations that Hersvall and Martinsson found in their investigations. The explanation of this is probably to be found in the fact pointed out earlier that the internationally evaluated compositions are in som cases extremes.

One might suppose that a combination of language measures would increase the correlation with marks. To examine this we recounted the values for the number of words and long words into standard points, and the means for these language measures were estimated. The correlation between these and the marks was .80. The correlation between marks and a combination of language measures therefore does not seem to be stronger as regards these compositions than if only the length of the composition, calculated in the number of words, is used.

The number of words was estimated also in the compositions in the benchmark sets for the written tasks numbers 1A, 1B and 1C. They are 10 up to 12 per task. Even in the case of these shorter compositions the correlations between the international mark and the number of words were high which is shown in table 36.

Table 36. Correlations between international marks and number of words in tasks 1A, 1B and 1 C

Written tasks		Correlation
1A	grade 6	.68
1A	grade 9	.90
1B	grade 6	.84
1B	grade 9	.75
1C	grades 6+9	.84

The mean of the five correlations is .80. These results attained with written tasks of different kinds provided a stimulus for a more comprehensive study. Out of the material from the main investigation were chosen at random 100 compositions per grade in task 6. The number of words in these 300 compositions was counted.

The number of words indicates a distribution from 12 up to 741 words. As can be seen in table 37 the mean is lowest for grade 6 and highest for the upper secondary school students. Even this is an indication of the importance of the number of words.

Table 37. Frequencies, means and distribution of number of words in 100 compositions in each grade 6, grade 9 and grades 2^2, 3^3

Number of words	Grade 6	Grade 9	Grades 2^2, 3^3
12 - 50	4	-	5
51 - 100	35	5	5
101 - 150	19	13	8
151 - 200	16	13	9
201 - 250	11	12	9
251 - 300	8	15	12
301 - 350	6	16	13
351 - 400	-	9	12
401 - 450	1	8	12
451 - 500	-	4	5
501 - 550	-	2	4
551 - 600	-	-	1
601 - 650	-	1	1
651 - 700	-	2	1
701 - 741	-	-	3
M	153	279	306
s	86	131	161

Some comparisons can be made with calculations from other investigations. Björnsson has collected compositions in the compulsory school that treat different topics and were without time limitations. The results were reported by Åke Pettersson (1989). In grade 6 the average number of words was 170 and in grade 9 260. The values are near what the students achieved in the IEA study, which however referred to a topic which was difficult for pupils in the nine-year compulsory school and in addition to that their writing time was limited.

The pupils generally write longer compositions when they relate or use their imagination than if they have to analyse, compare or argue.

A comparison can also be made with the pupils' productivity in the standardized achievement tests in grade 9. The number of words has been calculated in the compositions that during the years 1984-88 were included in the evaluation material, i.e. the compositions that are standardized and commented on centrally as a guidance for the teachers in the schools. In table

38 the average number of words per composition has been given. In addition to that a distinction has been made between compositions that are narrative or nearly only narrative, e.g. "Memories from a journey", "Teenage life", "Sports and I". These have been put in group I. The topics in group II are more demanding. According to the test constructors the pupils are expected to argue and make comparisons. Such topics are for example "Young in Sweden and in some other country", "Young then and now", "Sports on equal terms", "A place in my heart".

Table 38. Average number of words and means for compositions in the standardized achievement test during the years 1984-88

Number of compositions	Average number of words			Means	
	All	Group I	Group II	I	II
50	427	436	406	2.94	3.64

The mean number of words in group II is 406, and that value should be compared with the mean value 279 for grade 9 according to table 37. That the pupils in the standardized achievement tests write longer compositions is due partly to the fact that they have considerably longer writing time - 120 minutes compared with 60 minutes in the IEA study - partly to the fact that they write a test which they know will influence their final mark. It also seems to be a fact that the compositions in group II represent mainly better compositions to judge from the mean. Probably the average number of words would go down if all pupils had written the argumentative essay.

The compositions in task 6 had been scored by three teachers. The mean for the overall impression was correlated with the number of words in the compositions. In table 39 correlation coefficients are reported.

Table 39. Correlations between marks and number of words in compositions in task 6

Grade	Correlation
6	.58
9	.67
2^2, 3^3	.76

The correlations between marks and number of words are lower than in the calculations that were made regarding the compositions in the benchmark set. The correlation for these was .80 according to tables 35 and 36. Lower correlations can however be expected in the main investigation, as it comprises compositions with a normal distribution of marks.

It is obvious that the correlation is lowest in grade 6 and increases with higher grades. A difference like that is also what could be expected. In the discussion in "Correlation between the scorers" it was stated, that compressed scoring scales for the grades is a consequence of a common scale being used. Such a compression is strongest in task 6, which was written by three grades. A measure of to what extent the scoring scale was used is given by the means and the standard deviations in table 17. In the case of grade 6 the mean of 2.1 and the standard deviation of .7 mean that the main part of the pupils have received twos and only a small proportion of them fours or fives. The scale for the upper secondary school students has been moved upwards, but the mean 3.3 and the standard deviation of 1.0 indicate that also the lower marks are relatively frequent.

As a scale with fewer scoring steps gives lower correlations than a multigraded one, it is natural that especially grade 6 has a lower correlation. In this grade mainly the marks 1-3 have been used.

The correlations between marks and number of words in grades 9 and 2^2, 3^3 are therefore probably more reliable values. One can suppose that the correlations would be even higher, if the calculations had been made for a task in which a grade had been estimated according to their own scale.

It is however possible technically to correct the values for grade 6, as they have been calculated on a selected sample within the total distribution. In order to determine the corrected correlation coefficient we have estimated the total standard deviation for the three grades according to Garrett (1958, pp 55-56). The correction of the correlations has been made according to Thorndike's formula (1949, pp 169-176), and it gives for grade 6 the correlation .71. The correlation between marks and number of words is thus for grade 6 after correction of the same scope as in grades 9 and 2^2, 3^3. The value for grade 9 is raised after correction to .73, whereas the correlation for grades 2^2, 3^3 remains unchanged .76.

The correlations between the means and number of words turn out to be equally high as those between the raters or even a little higher. These are between .63 and .75 according to table 33. One may therefore draw the conclusion that a calculation of the number of words is an equally certain measure in an evaluation of compositions of this kind as using teachers as judges.

Of course it is not surprising that longer compositions should generally be considered better than shorter ones. It is quite natural that the more ideas, opinions and arguments that a pupil can bring up the more words are needed. The number of words becomes a measure of quality of content. Hultman and Westman have discussed this question and mean that the number of words can "be interpreted as a natural and good measure of a pupil's linguistic productivity". By that they mean not only "the simple ability to pile up words in heaps as big as possible. Rather it is a symptom of a linguistic proficiency which manifests itself in showing its ability to solve a problem" (Hultman, 1977, p 55).

6 Relations between background factors and the students' writing capacity

6.1 Factors that influence the students' writing capacity - a theoretic model

The writing results are influenced, like all other school performances, by many different conditions. Some of them have a direct influence on what the pupils produce and others have a more indirect effect. During the first years of the project much attention was paid to the designing of a theoretic model that would be able to describe factors that influence the pupils' performances and which might also be the basis of continued work. Not least important was having such a model to start from during the designing of tasks and questionnaires.

In the international comparisons it was of course necessary to observe the national background, which is of importance to both quantitative and qualitative differences in the pupils' writing. It is to be expected that what is taught in different countries is dependent among other things on the national linguistic situation and the literary background. The shape of the educational system also plays an important role. In a centralized educational system which lays much stress on official examinations one may presume great similarities of teaching among the schools. One can assume that teachers and pupils are well agreed on the content of their teaching programme. In such a system there is less space for local variations in the design and content of teaching.

The less centralized the school system of a country is the greater the possibilities of the individual schools are to create their own profiles. Each school can develop according to its own conditions and purposes. Of great importance to the development of the individual school is however such conditions as the socio-economic and linguistic surroundings of the school. Also the adults' attitude towards the individual school's aims and ambitions is important just like their willingness and readiness to identify themselves with and take part in the work going on at school.

It can be assumed that school organisation measures have an influence on the pupils' opportunities of performing to the best of their abilities. The conditions are different in schools with or without differentiation, in schools of different sizes and with different numbers of pupils in the classes. But also in schools, where the cultural atmosphere and school organization conditions seem to be equal, pupils in different classes can have varied possibilities of learning to write. For it may be assumed that well qualified teachers can give more stimulating teaching and supply the pupils with better knowledge and skills than unqualified teachers. An experienced teacher can probably plan his teaching in such a way that it gets a distinct structure, he can use different methods to stimulate the pupils and apply praise and criticism by way of motivation.

The influence on the pupils' writing achievements that may be referred to the teacher are due not only to the teacher's education and experience. It is also probable that for example the time set aside by the teacher for the teaching of writing is of significance, just like the teacher's interpretation of the aim of his teaching and thereby also the emphasis on writing skill, choice of teaching methods, frequency of exercises as well as ways of revision and feedback.

It has been proved in many investigations that the home is of importance to the pupil's success at school. Distinct agreement has been found for example between family background and interest in school activities on the one hand and the pupil's achievements on the other. It can be assumed that such factors also are significant to the pupil's writing performance. In addition to that the linguistic background of both the pupils and their parents is of importance. The effects of what language or languages are spoken at home are in all likelihood greater on writing than on any other school subjects.

The pupils' possibilities of performing well in writing are dependent on their opportunities of learning to write. These are not exclusively depending on the design of teaching. Of essential significance is their motivation to engage themselves in the teaching programme. Their progress is dependent on how necessary they think it is to be able to write, how important it is for their future, how much they need to practice and make an effort, how they usually succeed etc. To what extent they can benefit by their teaching is of course also dependent on their intellectual capacity and their interests outside school in reading and writing.

These discussioons have been made concrete by Takala (1988) in a theoretic model, which is reproduced in figure 5. The model comprises such factors as have been discussed above, and it endeavours to indicate not only their relations with the dependent variable written achievements but also the correlations between the factors. The arrows in the thought construction symbolize hypothetic causal relations.

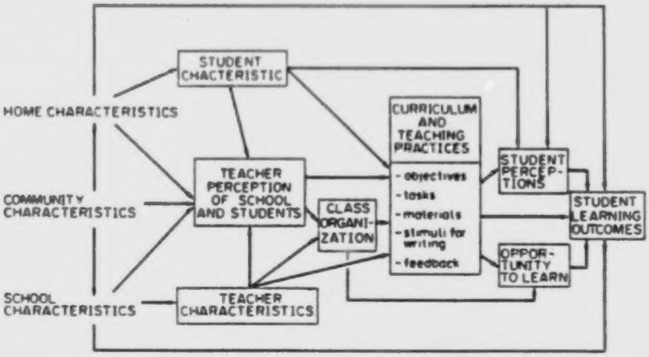

STUDENT CHACTERISTIC
HOME CHARACTERISTICS
COMMUNITY CHARACTERISTICS
SCHOOL CHARACTERISTICS
TEACHER PERCEPTION OF SCHOOL AND STUDENTS
TEACHER CHARACTERISTICS
CLASS ORGANI-ZATION
CURRICULUM AND TEACHING PRACTICES
- objectives
- tasks
- materials
- stimuli for writing
- feedback
STUDENT PERCEP-TIONS
OPPOR-TUNITY TO LEARN
STUDENT LEARNING OUTCOMES

Figure 5. A theoretic model of factors with importance to pupils' writing capacity

6.2 Structuring of information

The information about the pupils and their home conditions, on teaching of writing and on their school has been collected by means of the questionnaires described in paragraph 2.3.3. They contain a lot of questions, of which some are divided into partial questions. The school questionnaire comprises 150 variables, the teacher questionnaire about 450 and the pupil questionnaires around 100. This information must be arranged one way or another to be handy. For a detailed description of the dimension analyses see Löfqvist (1989e).

By means of factor analysis one can structure material by extracting factors that collect a number of related variables. The factor analyses of the questionnaires have been made with the use of the principal component method. Kaiser's criterion has been used to determine the number of extracted factors.

The questionnaires contain a number of questions that have not been included in the analyses. This has reference above all to questions whose answers cannot easily be ranked and questions with very skewed distributions of answers.

The factor analyses of the questionnaires led to a concentration of the variables in the school leader questionnaire down to 6 factors, in the teacher questionnaire down to 29 and in the pupil questionnaire also 29. The school leader questionnaire had been analysed in full, whereas the others had been divided into sections, as they contained too many variables to be treated in just one factor analysis. The teacher questionnaire was divided into eight sections.

Section	Number of variables
Education and experience	30
School and class	21
Important aspects of subject	20
Aims and working methods	15
Teaching techniques	20
Writing tasks	30
Exercises directed towards writing	16
Feedback	24

The pupil questionnaires had been divided into five sections.

Section	Number of variables
Education and linguistic background	26
Mass media and books	16
Cooperation within the family	23
Compositions and tests	7
Writing at school and in leisure hours	43

As the number of factors was still too large for continued analyses, a reduction of the number was tested by means of an analysis of the teacher questionnaire at a second level. In this analysis a large number of factors came out as regards the overall material, but for the separate age groups it was impossible to arrive at a solution.

The number of factors was therefore reduced on a logical basis. Those that were considered to be able to forecast writing skill and define the content of the model in figure 5 were chosen, which in all gave 25 factors. As even this meant too many to be included in a causal analysis, they were all correlated with marks. For well over half of them the correlations were very low and they could therefore be excluded.

For a preliminary causal analysis ten factors were selected. In the following survey the letters in front of the factor number indicates that the factor was established from the headmaster questionnaire (H), the teacher questionnaire (T) or the student questionnaire (S). In addition to the factors the variable of sex of the pupil was included. As dependent variable the marks in task 6 were used.

S	Sex
S9	Active intellectual spare time
S12	Swedish important according to the home
S3	Educational aspirations
S4	Watching TV
S29	Attitudes towards writing
T2	Own education in Swedish
T5	Teaching time of Swedish in school
T7	Grammar, important aspect of the subject
T14	Ordinary prose as stimulus material

As can be seen in the survey above the factors from the school leader quetionnaire had disappeared. Their correlations with marks were very low. The factors S9 and S12 describe the home conditions according to the students, T2 gives teacher characteristics and T5 the time for composition writing with preparation and revision included, T7 and T14 describe teaching and S3, S4 and S29 supply pupil characteristics.

The preliminary analysis comprised 376 pupils in grade 9. The solution indicated almost exclusively strong relations between factors that had been taken from the pupil questionnaires and the marks. Of importance to the pupils' achievements is cooperation within the family, that is what is defined by the manifest variables in factor S9 Active intellectual spare time. In that factor are included for example mutual discussions on books, radio and TV programmes, debates on family affairs and common events, conversations on work and school. The girls have a more positive attitude towards writing and also get higher marks for their compositions. Those who watch TV, however, do not succeed so well, and this seems to apply to boys for the most part.

The factors from the teacher questionnaire, however, turned out not to have any relations with marks. The different factors - the teacher's educational background, time of education in Swedish, teaching of grammar etc. - were related to each other but had no connection with the pupil's achievements. As it was considered urgent to try and find teaching variables of importance to the pupils' written achievements, the preliminary analyses were continued.

Some of the questions in the teacher questionnaire are built on variables that are meant to define different ways of writing. With the help of one of the constructors of the teacher questionnaire I had chosen in another connection a number of variables for a factor analysis of the total international material from all countries in the IEA study. They are represented by questions, in which the teachers are asked to state, to what extent they have emphasized different parts or skills in their teaching of their mother tongue or applied different procedures or made use of special exercises.

The factor analysis is thus not based only on the answers from the Swedish teachers but on the total number of teachers that are included in the IEA study in the age group that corresponds to grade 9 in Sweden, in all 1118 teachers.

In table 40 the chosen variables with their numbers from the teacher questionnaire are reproduced. The extracted factor loadings are indicated with two decimals without a whole-number zero.

Table 40. Factor matrix for 1118 teachers within the total international material

Variable no.	Description	1	2	3	4
V 330	Structures of different types of text (e.g. narration, exposition)	00	31	45	05
V 350	Communication theory (e.g. communication situation, speaker-audience relationships, etc.)	-01	-10	55	01
V 355	Aesthetics and literary theory (e.g. including genre theory, criticism)	02	-02	49	-02
V 440	Poetry	02	20	67	-01
V 445	Drama	-05	18	63	06
V 450	Prose fiction	01	34	67	-05
V 520	Reading	02	-41	36	02
V 565	In my class writing is an academic activity: students learn to write on school subjects for academic purposes.	27	67	01	09
V 567	My students write on an unprepared topic, without gathering extra information from various sources.	37	53	-40	08
V 568	In my teaching I devote attention to the stages of the writing process (planning, writing, editing).	41	42	-12	27
V 570	I prepare my students by discussing the topic with the class and by collecting words and phrases which are related to the topic.	29	81	-31	20
V 573	My students work in small groups. They write a joint composition.	30	77	-46	17
V 574	My pupils prepare their compositions in group-work. They work together while discussing the topic, outlining and revising, but they write their own essay.	30	81	-40	18
V 576	In my class the students first gather information on the topic they are going to write about.	19	52	-14	11
V 577	I prepare my students by discussing the topic and making an outline in class, before they start writing.	29	75	-40	19
V 578	I my writing instruction I let the students discuss the topic informally before writing, in order to collect and generate information and ideas on the topic.	30	71	-43	15

Table 40. Factor matrix for 1118 teachers within the total international material (cont.)

Variable no.	Description	1	2	3	4
V 634	Selected essays written by other pupils (somewhere else in the world) for analysis in class.	82	22	06	23
V 635	Selected essays written by professional authors to be studied as examples (rhetoric texts).	76	21	06	25
V 636	A collection of reading on a variety of subjects for analysis.	73	27	-02	25
V 637	A selection of compositions written by the students themselves used as material for analysis.	80	21	-02	21
V 638	A text (book) or instruction with general guidelines or prescriptions for writing a composition.	80	21	-06	22
V 639	Published criteria for good writing; e.g. official examination criteria, scoring guides, scored sample essays.	80	21	-10	16
V 723	Exercises in the organization of paragraphs and whole compositions.	52	14	08	65
V 724	Exercises in the development of organization: how to systematically treat a subject.	49	19	-04	72
V 725	Exercises in correcting and synthesizing materials from (reference) books and other sources.	48	18	04	73
V 726	Exercises in the development of context: thinking about the topic and developing the composition ("heuristics")	48	18	04	73

Four factors have been extracted. Factor 1 gathers the variables which are oriented towards the product, the complete composition. They describe the teachers' methods of giving the pupils patterns or models for what a good composition should look like. The teachers can analyse pupils' compositions and texts written by professional writers together with the pupils and also discuss general outlines and criteria for evaluation. Different exercises in collecting material, in organizing this and in how to treat the composition topic systematically have essential loadings as well on this factor. It can practically be characterized as a product oriented method designed for the teaching of writing.

Factor 2 is clearly oriented towards preparation and support of the pupils for their written tasks. Above all variables which can be said to characterize a process oriented type of writing have high loadings. Such are for instance the different possibilities of preparing the individual composition: to discuss the topic with the pupils, to help them with words and expressions, to outline the content or let them themselves give each other ideas and points of view before they start writing. Some variables seem to have loadings that are too high, among others the instruction about letting the pupils write about unprepared subjects. That variable is more product oriented. A name for the factor might be process oriented writing.

Som teachers seem especially to emphasize literature and reading-oriented aspects in their teaching. As can be seen in the factor matrix all such variables load very positively on factor 3, whereas the process-oriented variables have negative or insignificant loadings. It would be appropriate to call this factor reading-oriented teaching.

Mainly those variables that describe different exercises in how to organize the content of a composition build up factor 4. They rather characterize a methodology of formal writing exercises, but the variables also load fairly heavily on factor 1.

The factors can be said to define above all three types of teaching the mother tongue. Teachers may either be oriented mainly towards the complete composition and its design or else they can especially take an interest in work which is meant to lead up to the written product, i.e. preparations and the very process of writing. The third teaching method is above all characterized by the focusing on the reading of fiction and a fourth method maybe by formal exercises.

The variables in table 40 were correlated with the pupils' written results in the Swedish investigation. They however showed only low correlations. There was on the whole, either in the school leader or teacher questionnaires, any variables with a relevance to the teaching of writing that had any correlation worth mentioning with the pupils' composition marks. The variables that the continued analyses therefore came to be based on were only taken from the pupil questionnaires.

In the pupil questionnaire there are also som questions referring to the number of compositions of different length that the pupils say they have to write. One might hope that a variable of that sort would correlate a little stronger with marks, but even for these the correlations were weak, between .08 and -.08.

In the student questionnaire there is a question with descriptions of different written activities, and the pupils are to indicate how often they have been given such tasks. The correlations between the frequencies of the tasks and the written performances (task 6) are shown in table 41. The descriptions have been shortened.

Table 41. Correlations between written activities according to the pupil questionnaire and written performances

Variable no.	Description	Correlations		
		grade 6	grade 9	grades 22, 33
*V 264	Personal experience	.11	.10	.01
V 265	Event: witness of	.03	-.02	.05
V 266	Writing for different readers	-.04	-.04	-.06
V 267	Telling about how different persons see it	.02	-.09	.03
V 268	Note while you are listening	-.04	.07	.26
*V 269	Summing up a book	.01	.13	.09
V 270	Summing up a discussion	.10	.03	.10
V 271	Summing up a lecture	.04	.11	.14
*V 272	Making up a story	.10	.13	-.06
V 273	Writing a poem	.04	.07	-.03
V 274	Writing a conversation	.12	.18	.00
V 275	Writing down one's thoughts	.04	.11	.19
*V 276	Writing one's own opinions	.10	.08	.16
*V 277	Reviewing a book	.09	.18	.10
V 278	Impersonal letter	.00	-.02	.00
V 279	Practical instructions	.00	.06	-.01
V 280	Messages	.10	.00	-.12
V 320	Describing an object	.08	.01	.02
V 321	Describing something happening	.03	.11	.20
V 322	Describing a person	.11	.06	.10
V323	Describing a problem	.03	.05	.09

The correlations are on the whole low. Some tasks have however been chosen to be included in the continued analysis. They have been marked with an asterisk in the table.

The reason for the choice of the activities marked with an asterisk is that teachers and pupils say they often occur during teaching hours (Löfqvist, 1988b), that they represent different types of writing and that they almost entirely have positive correlations, even if these are low.

Some further variables ought to be interesting. In a factor, Books at home, from the pupil questionnaire the variable "Reading, not cartoons" was included. It describes leisure time reading, cartoons not included and correlated between .11 and .17 with written results. It was selected not only guided by the numerical connection but above all because an interest in reading ought reasonably to be related to writing skill.

The questions on the parents' educational background were not possible to use, as many pupils had not answered them. Even the answers the question on the length of the parents' schooling and education were very uncertain. On the other hand the pupils were relatively good at telling about mother's or father's occupation (Löfqvist, 1988d). The occupations have been ranked according to a socio-economic classification which Statistics Sweden uses. These facts about the parents' occupations are the best measure of the parents' education and have therefore been used even if they turned out to have relatively weak correlations with the pupils' achievements.

The latent variable WRITE in table 42 consists of different writing tasks that the pupils get at school. What they write about during their spare time can be found in ATTITUDE as the voluntary writing of diaries and letters is a manifestation of their attitude towards writing.

The definite selection of variables, which all of them are to be found in the pupil questionnaires, are shown in table 42. As a measure of the reliabilities Cronbach's alpha has been used. The abbreviations of the latent variables have been used at the report on results in the next chapter.

In HOMECULT some related manifest variables have been added to build up H1, H2 and H3 (see table 42). At the examination of reliability it turned out to be lower in the upper secondary school because of disturbance from H2, which is probably due to the fact that the upper secondary school students do not talk about school and what happens there to the same extent as compulsory school pupils do. H2 was therefore removed as regards the upper secondary school. In other respects the same set of variables applies to all populations.

Table 42. Survey of variables for causal analysis

Latent variables	Manifest variables	Direction	Reliability coefficient grade 6	Reliability coefficient grade 9	Reliability coefficient grades 2^2, 3^3
SEX	pupil's sex	boy-girl	-	-	-
PAR OCC	mother's occupation	low-high	.61	.53	.56
	father's occupation	low-high			
HOME CULT	discuss books + discuss TV (H1)	seldom-often	.70	.70	.63
	tell about school + talk about job and school (H2)	seldom-often			
	disc family affairs + disc general affairs (H3)	seldom-often			
READ FOR FUN	reading (not cartoons) for fun	not at all-many hours	-	-	-
CARTOONS	reading cartoons	not at all-many hours	-	-	-
TV	TV during schoolday	not at all-	.60	.69	.65
	TV during holiday	many hours			
WRITE AT SCHOOL	personal experience	never-often	.46	.57	.54
	summing a book	never-often			
	make up a story	never-often			
	one's own opinion	never-often			
	reviewing a book	never-often			
ATTITUDE TO WRITING	attitude to writing	boring-fun	.54	.58	.58
	diary in spare time	no-yes			
	correspondence	no-yes			
	comfort at school	bad-good			
EDUC ASPIR	planned studies	number of years	-	-	-
MARKS	overall impression	low-high	.95	.96	.94
	content and disposition	low-high			
	style and linguistic form	low-high			

BEST COPY AVAILABLE

6.3 Partial Least Squares (PLS) - a short presentation

The analysis technique which has been used to examine how different factors are related to each other is called PLS (Partial Least Squares) and was developed by Wold (1982) and Lohmöller (1981). The technique implies that from a theory a model is built up showing how different factors are causally related to each other. In a so-called measurement model the relation of the manifest variables to their respective factors, i.e. the latent variables are specified. The relation of the latent variables to each other is specified in a so-called structure model. The theoretically-based model can now be tested by means of collected data. One then gets information on possible deficiencies in measurement model as well as structure model. If the model is adjusted to data, the result one receives should be tested on new sets of data (so-called cross-validation).

Somewhat simplified one could say that the PLS technique at the same time makes use of both factor analysis and regression analysis. The structure analysis is thus a path analysis with latent variables, where one tries to find the maximal prediction from independent variables to the dependent ones. To use latent variables means a more correct estimation than regression analysis with manifest variables. In regression analysis one gets an overestimation of the relations, especially if one works with many variables compared with the number of individuals.

Developing a causal model with the help of such a technique naturally does not mean that one has found a proof of causality. We can however say that we have obtained empiric support, i.e. data are compatible with a causal interpretation according to the model. PLS is a powerful method for the study of complicated relations between many variables, especially in situations in which it is impossible to make use of experimental procedures.

6.4 Relations within background factors and the written achievements within different age groups

The analyses in which the PLS method was used to test the theoretic model and measure how strongly different factors result in good written performances have been made in all three age groups. As dependent variable in grade 6 the overall impression for task 5 has been used, the personal story, and the calculations are based on 850 pupils with complete data, in grade 9 task 9, advisory letters, has been used, with 1123 pupils and in grades 2^2, 3^3 task 7, the reflective essay, with 294 pupils. The total number of pupils who wrote the compositions in each grade was 1187, 1606 and 676. The dropping off is due to the fact that certain pupils delivered blank or illegible answers and that some pupils misunderstood the task and did not get any marks, but the main reason is that pupils do not have values for some of the variables that are included in the analysis. The results thus refer to those pupils that have complete data.

In other connection we have observed that the pupils in this IEA study say that they get on well at school (Löfqvist, 1989d). Of the intermediate stage pupils 80% find schoolwork amusing or very amusing, and at the senior stage and in the upper secondary school the share is well over 60% of the pupils. Only few pupils say that they do not feel comfortable in school. Only a few percent find it boring or very boring, whereas the others find it difficult to take up a definite position. The attitude towards writing is highly positive, even if the values are not as high as for comfort and well-being at school (Löfqvist, 1988 b). Its shows in figures 6-8 that the positive attitudes are mainly connected with the girls. The direct relations from sex to attitudes are very strong with .58 in grade 6, .63 in grade 9 and .61 in grades 2^2, 3^3.

The girls are not only more interested in writing that the boys. They write more often during their spare time with personal stories, diary and letters as their most popular writing acitivities. Their performances are also better than those of the boys. The direct relation from sex to marks is .18, .20 and .28 in the different age groups. We thus find that the difference of achievement increases between girls and boys from lower to higher ages. In table 43 can be seen not only the direct but also the total relation from each factor up to the result in the form of marks.

It has earlier been maintained that the girls get better results in writing. In an examination of linguistic differences between sex, home and schools it was stated that the girls' compositions in grade 6 were scored as definitely better than those of the boys (Lundquist, 1978). Their means have in most cases turned out to be higher than those of the boys in the standardized achievement tests in writing in grade 9 (Eneskär, 1988). The language of the upper secondary school students has been examined by Hultman and Westman within the project "Writing syntax" (1977). In a random selection of compositions from the standardized achievement tests the girls received higher marks than the boys by independent raters. Larsson (1984) writes that the difference in written performance between the girls and the boys in his material is considerable in the girls' favour.

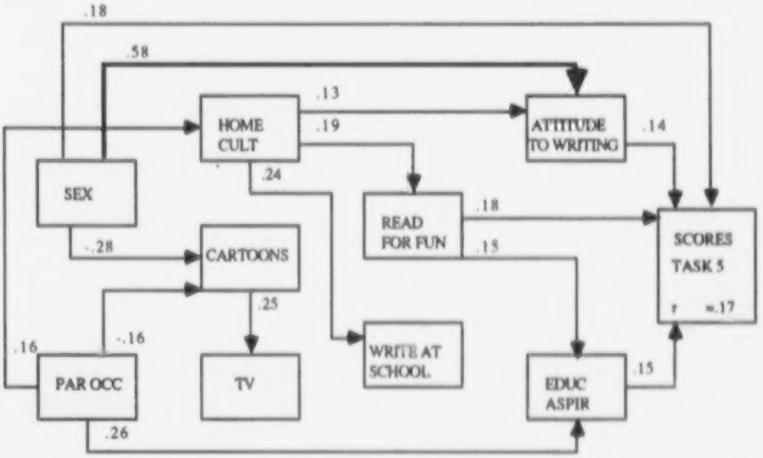

.18
.58
.13
.19
HOME CULT
ATTITUDE TO WRITING
.14
SEX
.24
READ FOR FUN
.18
.15
SCORES TASK 5
r =.17
-.28
CARTOONS
.25
-.16
.16
WRITE AT SCHOOL
PAR OCC
TV
EDUC ASPIR
.15
.26

Figure 6. Results of PLS analysis in grade 6 (n = 850)

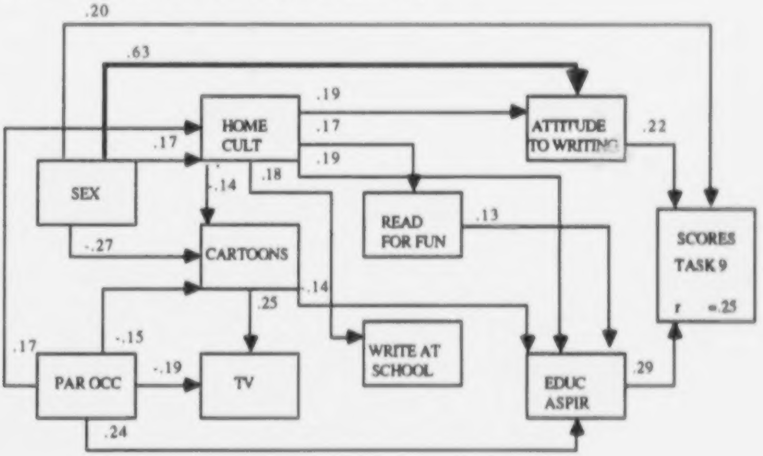

.20
.63
.19
HOME CULT
.17
ATTITUDE TO WRITING
.22
.17
.19
SEX
.14
.18
READ FOR FUN
.13
SCORES TASK 9
r =.25
-.27
CARTOONS
.14
.25
-.15
.17
WRITE AT SCHOOL
PAR OCC
-.19
TV
EDUC ASPIR
.29
.24

Figure 7. Results of PLS analysis in grade 9 (n = 1123)

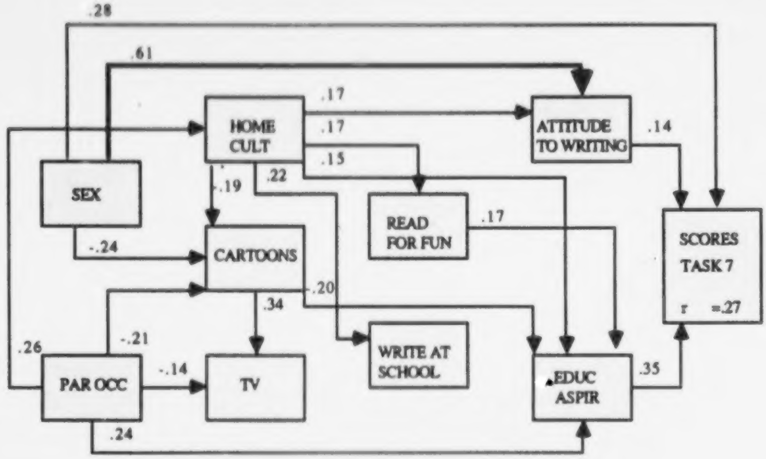


Figure 8. Results of PLS analysis in grades 2^2, 3^3 (n = 294)

In a study of segregation in the compulsory school Arnman and Jönsson (1983) find that the girls manage better at school. They get better results in the standardized achievement tests and higher marks than the boys in grade 9 in Swedish and matemathics. In addition to that the pupils' achievements are clearly related to the parents' social group. High-status classes in most cases get better results at school than low-status classes. This correlation between the performances at school of pupils from socially and culturally different homes has been established in many studies (Grundin, 1975; Hansson, 1975; Grogarn, 1979).

Hansson (1975) states that the social background is the most essential explantion of differences between pupils' performances not only in the subject of Swedish and not exclusively in Sweden. The same result has been achieved in Hansson's studies of pupils' reading ability and understanding of literature and also in other subjects, e.g. science. The results have likewise been the same in other countries in these IEA studies. Schoolwork is naturally important to the development of the pupils but "what the result shows is that school does not manage to remove those differences of knowledge and skills that are due to the fact that the pupils bring with them different quantities of those from their home environment when they first begin at school." (p 175)

The students in Arnman and Jönsson's study wrote compositions on their visions of the future in grade 8. Nearly all of those in social group 1 intended to continue in the upper secondary school and of these 83% planned to begin on 3- or 4-year theoretic lines. The majority of the pupils from social group III also meant to apply for the upper secondary school, but 62% of these mentioned 2-year vocational or semi-theoretic lines.

In this IEA study we find in all pupil groups a relation from the parents' occupation to the pupils' educational aspirations and indirectly to their written achievements. The higher the parent's occupational status the longer education the pupils intend to have. The correlation between the parents' educational background and the pupils' attitude towards continued studies has also been pointed out in pervious studies (Boalt, 1947; Härnqvist & Grahm, 1963; Boalt & Husén, 1967).

Jönsson (1985) observes in her survey of literature on TV consumption and social background that a correlation between the children's TV watching on the one hand and the family's educational level and social group status on the other has been found in many studies. Children in families, where the parents have higher education and higher socio-economic position watch TV less than others. The results in Swedish investigations, among others in Jönsson's longitudinal study of the relation between TV and school, is confirmed by foreign studies. It is therefore not surprising that the relations in the IEA study from the parents' occupations to the consumption of cartoons and TV should be negative. In grade 6 there is no direct relation with TV watching but indirectly by way of cartoon reading the total effect is negative. The relation was to be expected.

The relation between cartoon reading and TV watching is conspicuous. He who reads a lot of cartoons also watches many TV programmes. That relation is strongest in the upper secondary school. We furthermore find another difference between the age groups. In the case of the older pupils there are stronger negative relations from the parents' occupation to cartoons and TV than in the intermediate stage pupils. It may be due to the fact that the pupils in grade 6 are less influenced by their parents' attitudes in their media consumption than the older pupils. Among these there are greater differences in their TV watching, which may have its basis in the attitude of the homes towards cartoons and TV. Home culture is negatively related to cartoons in grade 9 and in grade 2^2, 3^3 but not in grade 6. According to the pupils' information it is mainly those in grade 6 that read comic papers (Löfqvist, 1989d). About 2-4 hours a week are used by two out of three pupils on reading of that kind. With increasing age the interest diminishes, and half of the upper secondary students do not read comics at all. The intermediate level pupils are also the most eager consumers of TV programmes. The difference between the senior level- and intermediate level-pupils is slight. The upper secondary school students, however, say that they spend considerably less time on TV watching. That the analysis does not indicate any relation from home culture to mass media consumption at the intermediate level may seem surprising and will be commented on later.

Larsson (1984) does not distinguish between age groups in his investigation but finds that written performances are generally a great deal lower in the case of pupils that spend much time on TV watching.

Boys read more comics than girls. They also watch TV more than girls, and there are many both Swedish and foreign proofs of this (Jönsson, 1985). We find the same results in this investigation with an indirect relation from sex over cartoon reading to TV watching.

Table 43. Direct and total relations to marks in grades 6, 9 and 2^2, 3^3

From factor	Grade 6		Grade 9		Grades2^2, 3^3	
	dir	tot	dir	tot	dir	tot
SEX	.18	.26	.20	.37	.28	.38
PAR OCC	.00	.05	.00	.09	.00	.13
HOME CULT	.00	.06	.00	.11	.00	.10
CARTOONS	.00	.00	.00	-.04	.00	-.07
TV	.00	.00	.00	.00	.00	.00
READ FOR FUN	.18	.21	.00	.04	.00	.06
WRITE AT SCHOOL	.00	.00	.00	.00	.00	.00
ATTITUDE TO WRITING	.14	.14	.22	.22	.14	.14
EDUC ASPIR	.15	.15	.29	.29	.35	.35

The importance of the homes to the individual's intellectual and then mainly verbal development has been emphasized by Bernstein (Bernstein & Henderson, 1969). On the basis of a comprehensive field survey of mothers they considered themselves to be able to distinguish different linguistic patterns connected with social groups. In one of them, where leaders of a business and civil servants were dominating, great importance was attached to language in order to teach the children ethical norms, obedience, independent thinking and emotional reactions. The second group mainly consisting of industrial workers was characterized by a linguistically more restrictive style. In this the dialogue between parents and children did not play such an important role as in the higher social group. The difference of linguistic setting is of significance to the linguistic and social development of the children in both groups as well as to their chances of feeling at home in school.

The factor HOME CULT is built up from variables that give us a measure of the solidarity and contact between the family members. It can illustrate differences of linguistic habits and measure, to what extent children and parents help each other at home and discuss with each other. This factor refers not only to conversations on TV programmes, religion, politics or other general affairs but also reflects the confidence and trust that parents show in their children by including them in discussions about family economy and circumstances that concern household and relatives. The difference of linguistic surroundings in different social groups that Bernstein found also shows in the IEA study. The relation from the parents' occupational status to home culture in the three grades is .16, .17 and .26.

The atmosphere of the home has an influence on several other factors. It is important to all pupils as regards their attitude towards school and writing and in the groups of older pupils their plans for continued education. That the relation between home culture and educational aspirations is not established in grade 6 is probably due to the fact that intermediate level pupils have rather vague ideas about their studies after the compulsory school. It was pointed out above that home culture does not seem to influence the pupils' reading of cartoons and TV watching in grade 6. As the education and occupational status of the parents are positively related to home culture, one should have a right to expect a negative relation even to mass media consumption. Jönsson (1985) found that mothers with higher education and social group affiliation have a more restrictive and supervising attitude towards the children's use of TV in the nursery school and in grades 3 and 5 in the compulsory school. The supervising function diminishes, however, when the children grow older. The relation that Jönsson found but which has not been established in the IEA study may be explained by the fact that the factor HOME CULT, from the point of view of content, is broader and more comprehensive than in Jönsson's investigation, in which the influence of the homes on the children consists of comments, discussions and elucidations of TV programmes to help their children.

In all grades the home influences the children's interest in reading. Reading comprises all types of reading that the pupils devote themselves to for fun with the exception of comics. Interest in reading does not only influence the pupils' disposition towards continued education. At the intermediate stage we also find a strong relation from reading to marks (.18), whereas the relations at the senior stage and in the upper secondary school are weak or indirect (.04 and .06 respectively), which is shown in table 43.

In a previous report it has been pointed out that the pupils with increasing age get more aware of the scope and direction of continuous education (Löfqvist, 1989d). It is natural that the pupils should know better how to estimate their qualifications for further studies and that they should be able to make more realistic evaluations in the upper secondary school than in the compulsory school. This increased awareness of one's personal capacity is illustrated in the relation from educational aspirations to marks, which grows stronger and stronger from the intermediate stage (.15) to the upper secondary school (.35).

It was expected that there would be a possibility of finding that pupils who more often get certain frequent writing tasks would make better written performances than pupils who write such tasks less often. However, no such relations, direct or indirect, from writing to marks have appeared. On the other hand we can observe a relation from home culture to writing in all grades, which may imply that pupils with good family conditions consider that they get more opportunities to write at school than others. It is however also possible that pupils who go to schools with good social surroundings really get possiblities of writing a larger number of compositions than pupils from socially deprived areas.

In the box for the dependent variable scores, the variance which the factors of the analysis predict is mentioned. In grade 6 it is 17%, in grade 9 25% and in grades 2^2, 3^3 27%. There is thus a considerable part of the variation which is not explained by the analysis but are caused by other factors.

The solution that the PLS analyses have given is in many respects parallel to what we have found in another analysis that referred to an international comparison (Löfgren & Löfqvist, 1989). The important independent variables were also there the pupils' background conditions.

If we compare the PLS solutions that we received by the hypothetic model in figure 5 paragraph 6.1, we find that it is verified in certain respects. We have thus found that the material confirms the belief that the pupils' home background is important to their results in writing. The parents' education and the atmosphere in the home play an important role concerning the children's interest in schoolwork, writing and continued studies. It has not, however, been possible to make use of such factors in the model as the ones that are based on the school leader or teacher questionnaires, as the variables from these questionnaires did not turn out to be predictive. Collected data therefore do not give us evidence that for example certain teaching methods might be better than others.

Naturally one cannon state that the pupils' written achievements are exclusively dependent on their sex or attitudes and independent of the efforts of school. School education in all likelihood plays a part to how the pupils' writing capacity develops. One might however interpret the outcome to imply that the pupils' achievements are rather independent of what teaching method the teacher puts into practice. The pupils would for example learn how to write almost as well whether the teachers mainly emphasize preparations before the composition is written or place the most essential part of their writing instructions after the writing sessions in the form of discussion and correction of the compositions. It would also be of less importance if the pupils were asked to write compositions often or only on a few occasions during the term.

Such a conclusion may seem likely and simple, but probably the explanations are more complicated. Instruction in writing is characterized and determined not only by how the teachers have described in the questionnaire the way they go about it. Their possibilities of motivating and arouse enthusiasm in the pupils may depend less on the method being used than on factors connected with their personalities. A methodical variant which seems to be successful with certain teachers may perhaps be less fortunate with others. Therefore it might be difficult to find causal relations between teaching variables and written achievements, as long as one cannot keep the psycho-social climate in the classrooms under control.

Emphasis on the importance of the psycho-social factors must however not be taken as a pretext for the view that the choice of teaching method would be unimportant. Plausible explanations of the relations that failed to appear between the descriptions of what goes on in the classroom and the written achievements have also been discussed in the report by Löfgren and Löfqvist (1989) mentioned before. In it it was pointed out that

1) questionnaire data collected on one single occasion do not give a correct description of the linguistic influence that the pupils are exposed to during a number of years,
2) class data mixed with individual data may lead to false conclusions,
3) teachers' descriptions of education in questionnaires may seem equal, but teaching in the classrooms need not be understood unambiguously by the pupils,
4) factual questions on teaching must be supplemented with questions on the psycho-social climate in the class to make it possible to find causal relations.

The PLS solutions which are described in this report, have been carried out with variables exclusively from the pupil questionnaires on a pupil basis. The second explanation is therefore not applicable. The others are however equally essential. More important still is perhaps that the differences of method are marginal compared with the fundamental conditions, i.e. time, general school quality, common equipment etc.

6.5 Relations within the same age group for different written tasks

The PLS solutions that have been presented up to now - the analyses in the previous section and the analysis meant for an international comparison (Löfgren & Löfqvist, 1989) - indicate the same important explanations of the pupils' written results, namely sex, attitudes, socio-cultural background and media habits. Even if one varies the set of independent variables as has been done in the two PLS solutions, the essential causal conditions are not hidden. This is the strong point of the PLS technique. The researcher is free on a logical basis to choose the variables he wishes to use and also to decide in what order they should be entered into the analysis. The PLS technique does not however allow recursive relations. But there are possibilities of establishing within the same analysis a varying number of relations from independent variables to dependent ones. More of fewer paths can be opened in the structure model. Weak relations in this way get an opportunity of appearing.

With the purpose of compare different PLS solutions within the same grade some supplementary analyses were carried out. In grade 9 the same variables were used, among other things for an analysis with task 9, the advisory letter, as well as task 6, the argumentative essay, as dependent variables. In these analyses the latent variable READ FOR FUN has been built up with two manifest variables. In addition to "reading (not cartoons) for fun" is the variable "library visits included". Both of them were included in a factor "Books and home" in the pupil questionnaires. Otherwise the set of variables is the same as has been reported in table 42. The reliablity coefficient for Read is .46 (task 9) and .47 (task 6) and for the others the same as in the table.

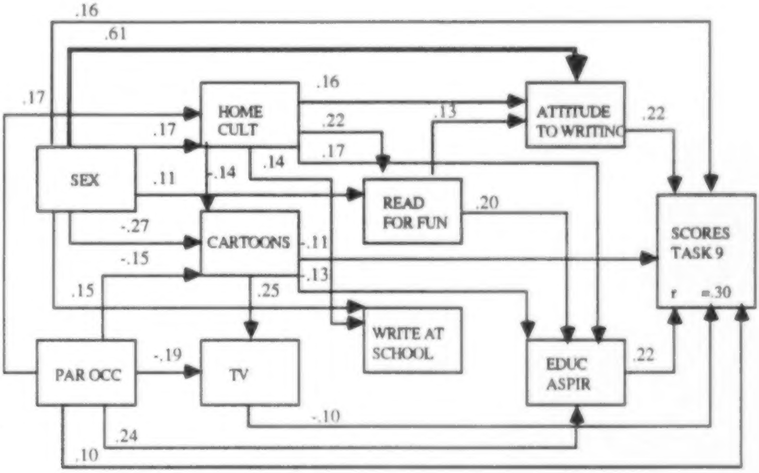

.16
.61
.17
HOME CULT
.16
.22
.13
ATTITUDE TO WRITING
.22
.17
.17
.14
SEX
.11
-.14
READ FOR FUN
.20
SCORES TASK 9
r = .30
-.27
CARTOONS
-.11
-.15
-.13
.15
.25
WRITE AT SCHOOL
PAR OCC
-.19
TV
EDUC ASPIR
.22
-.10
.24
.10

Figure 9. Results of the PLS analysis in grade 9 with task 9 as dependent variable (n = 1123)

The PLS solution for task 9 is shown in figure 9. It can be compared with figure 7, which also refers to task 9 with the same number of pupils and the same grade. In the solution in figure 9 several paths have been opened between the variables. More relations have been established, even if the majority of them are weak. We thus find relations from sex to read (.11) and to write (.15), from the parents' occupation directly to marks (.10), from cartoons to educational aspirations (-.13) and to marks (-.11), from TV to marks (-.10) and from read to attitudes (.13). Common for all is that they are between .10 and .15 and thus indicate a weak influence.

A comparison between direct and total effects can be made with the help of table 44. Sex, attitudes, educational aspirations and home culture have on the whole the same power. The parents' occupational status gets more accentuated as an explantion through the direct relation in figure 9, and this applies also to massmedia consumption. The variable WRITE is isolated in the two analyses and READ has a marginal influence on the written achievements.

Table 44. Direct and total relations in the PLS solutions in figures 7 and 9 to the dependent variable

From factor	Figure 7		Figure 9	
	dir	tot	dir	tot
SEX	.20	.37	.16	.37
PAR OCC	.00	.09	.10	.22
HOME CULT	.00	.11	.00	.11
CARTOONS	.00	-.04	-.11	-.16
TV	.00	.00	-.10	-.10
READ FOR FUN	.00	.04	.00	.07
WRITE AT SCHOOL	.00	.00	.00	.00
ATTITUDE TO WRITING	.22	.22	.22	.22
EDUC ASPIR	.29	.29	.22	.22

The PLS solution for task 6 is described in figure 10. Thus it concerns the same grade and the same variables as figure 9 but the dependent variable is different. A comparison can therefore be made to find out if different writing tasks within the same grade give the same explanations of the written performances.

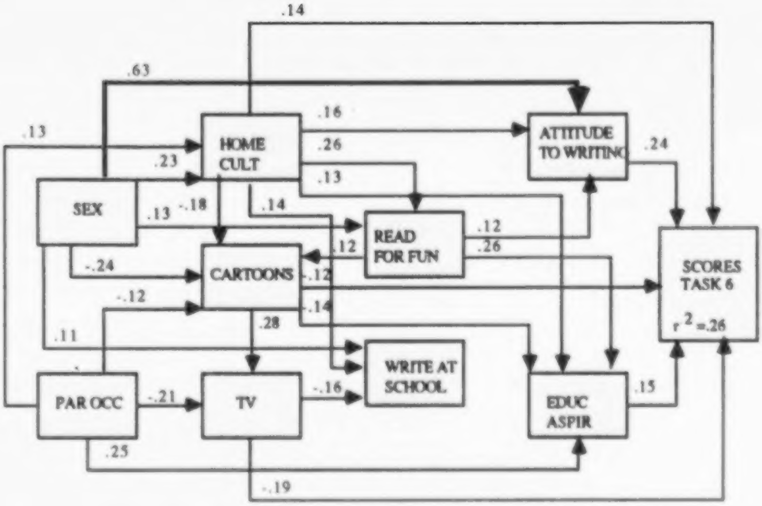

.14
.63
.13
.16
HOME CULT
.26
ATTITUDE TO WRITING
.24
.23
.13
SEX
.13
-.18
.14
READ FOR FUN
.12
.26
.12
CARTOONS
-.12
SCORES TASK 6
r 2 =.26
-.24
-.12
-.14
.28
.11
WRITE AT SCHOOL
PAR OCC
-.21
TV
-.16
EDUC ASPIR
.15
.25
-.19

Figure 10. Results of PLS analysis in grade 9 with task 7 as dependent variable (n = 387)

In figure 9 there is a direct influence from sex to marks (.16) and also from the parents' occupational status to marks (.10). These relations have not been established in figure 10, but instead there is a direct influence from home culture to marks (.14). Furthermore in figure 10 there is a relation from TV to writing (-.16). Otherwise the figures look alike with slight changes in the values.

We find certain variations between the solutions in the figures. An essential explanation of why the girls perform better in writing is their more positive attitude to school and their greater inclination to discuss school and family affairs with the other members of the family. There is an interplay between these variables and sex which makes the relation from sex stronger in figure 9, whereas home culture is more distinctly accentuated in figure 10. The mere fact that a pupil is a girl thus does not imply that she is able to write better than a boy. Cartoons and TV appear more distinctly in figure 10, but otherwise the two analyses with different writing tasks give almost the same result, which is also shown in table 45.

Table 45. Direct and total relations in PLS solutions in figures 9 and 10 to the dependent variable

From factor	Figure 9 dir	Figure 9 tot	Figure 10 dir	Figure 10 tot
SEX	.16	.37	.00	.26
PAR OCC	.10	.22	.00	.13
HOME CULT	.00	.11	.14	.24
CARTOONS	-.11	-.16	-.12	-.19
TV	-.10	-.10	-.19	-.19
READ FOR FUN	.00	.07	.00	.04
WRITE AT SCHOOL	.00	.00	.00	.00
ATTITUDE TO WRITING	.22	.22	.24	.24
EDUC ASPIR	.22	.22	.15	.15

It is important to observe that the explanation value in the analyses in figure 9 and 10 is somewhat higher than in the ones discussed earlier. In figure 9 30% of the variation for task 9 in grade 9 is explained, whereas the prediction of all the variation in the dependent variable essay scores is 25% in figure 7.

7 Contents of the compositions

The IEA studies are primarily comparative analyses. The purpose is to make international comparisons, among other things of teaching methods and pupil achievements. Also the study in writing has international comparisons as its aim.

With a direction like that it is natural that the methods of measurement should be the centre of observation. In this study a great deal of attention has been paid to the grading of the compositions. Much work has further been devoted to a description of the teaching method and an attempt to identify such factors that can explain variations in the pupils performances. Less attention has however been paid to the contents of the essays.

The choice of writing tasks as has been described earlier has been made very carefully so that these would cover different fields of the writing domain. This comprehensiveness together with the larger number of essays from different age groups makes the material a rich source for studies of the contents. Such studies would be possible to carry out to illustrate among other things how cultural differences between countries are reflected in the pupils' writing. It is probable that not only differences of content can be traced from the pupils' varying childhood and adolescence surroundings, literary traditions etc., but also stilistic differences, for example in the choice of style and linguistic tone.

In this thesis some notes have been made in different places on the content of the compositions. Paragraph 3.3.8, Students' world of ideas, is thus based on an analysis of the content of the argumentative task (Löfqvist, 1989b). The analysis is limited to a description of the pupils' choice of topic and of the differences in the choice of topics between the grades. It does not however refer to the way of arguing, the choice of reader etc.

In some tasks the pupils have been given a number of topics to choose from. In task 5 most pupils chose to write topics of imagination as has been described earlier in the paragraph "Scores for the written tasks". This result corresponds well with what the pupils have said that they preferred to write about (Löfqvist, 1988b). Irrespective of if they have written imaginative adventure stories or told about personal experiences, something that also belongs to their favourite topics, no studies have been made of the contents. The Swedish essays in one of the topics - My most frightening experience - are however used by a research student for a profound analysis.

The pragmatic writing tasks - description of a bicycle, messages to the headmaster and to the home, the self-description and the answer to an advertisement - are probably well suited for both international and national comparisons of contents. One can for example suppose that the pupils' choice of reason for not coming to the headmaster's office may vary a great deal just like the linguistic tone that the pupils use in their messages. Of the Swedish students some address themselves to the headmaster using a very familiar and informal tone, whereas others are ceremonial, especially in the greetings and complimentary closes. Such style differences are also very common in the answers to an advertisement. The differences in linguistic tone would probably seem striking in international comparisons of tasks of this kind.

Task 9, the advisory letter on writing, was meant to serve two purposes, namely partly to be an essay on how to write compositions, partly to inform on what pupils' in different school systems know and think about composition writing and scoring. An analysis of content has been made of the Swedish compositions, and the differences of opinion among the pupils in the compulsory school and the upper secondary school on what is important when compositions are evaluated has been accounted for in a partial report (Löfqvist, 1988b). The results are discussed in the following chapter 8.2, Reliablity of composition scoring. An international comparison of the results in nine countries is reported by Takala (1987). The pupils in grade 9 and corresponding grades in other countries except Thailand mean that correct linguistic style, spelling and punctuation are the most essential points. The young people from Thailand first of all emphasize organisation aspects, whereas linguistic formalities are very much accentuated especially in Chile and the Netherlands. Swedish pupils attach less importance to content than students in West Germany, Finland, Italy and New Zealand but stress more than others advice that refers to choice of topic, planning with the help of keywords and control of what has been written. The differences are however in many cases not great between the countries. No Swedish pupils mean that behaviour at school and personal neatness play a part for the scoring of compositions, but in other countries such advice is offered, even if only to a limited extent.

The investigation comprises material suitable for continued work, as can be seen from what has been hinted at here, but time and economic resources have not been sufficient for further contributions. In addition to analyses of contents the results furthermore give rise to new questions and problems which would be interesting to work at.

8 Pedagogical consequences

The IEA studies are characterized by their distinctly pragmatic character. The outcome of the usually arouse considerble attention, constitute an argument in the pedagogical debate and may bring about consequences for the ducational systems in the participating countries. Thus for example the publication of the results of the so-called six-subject investigation lead to a moderation in the very sharp attacks on the standard of attainment of Swedish schools during the first half of the 1970s. The Swedish pupils were not inferior to their companions in most of the comparable countries.

A new debate on pupil achievements flared up in Sweden in the middle of the 1980s, after the comparative results of the second investigation in mathematics had been disclosed. The Swedish pupils in grade 7 turned out to have got poor results, above all in algebra and arithmetic. The storm of criticism had a political impact and led to reinforcement of the teaching of mathematics at the intermediate level.

When the IEA project in writing started at the beginning of the 1980s one of the purposes of the study was to describe how students in different countries manage to accomplish the writing tasks. Just as earlier on, for example in science subjects and mathematics, it had been possible to measure the pupils' achievements and to make comparisons between the countries, the intention was to try and find an internationally accepted method for the evaluation of compositions in order to describe differences due to school systems and cultural variations in a more qualitative way. This aim has more and more been toned down, as little by little it has been certified the difficulties in establishing international scales that might be considered objectively satisfactory have not been possible to master (de Glopper, 1988).

8.1 Teaching of writing

One of the duties of the project has been to try and elucidate and describe the content of the teaching of writing, methodology and evaluation methods. A description has been made including comparisons with practice in other countries (Löfqvist, 1988c). The attempts to relate variants of educational methods to the pupils' written achievements has not led to any of that type appearing in this study, which has been shown in the discussion in chapter 6.

It can however be debated if the methods that Swedish teachers often use are suitable. Such a discussion can have its starting-point in methodological considerations but may also be based on comparisons with how in this domain work is carried out in other countries. Experiences abroad can give impulses to changes and be stimulating to one's own teaching.

In the chapter on teaching in writing there was a short reference to the teachers' descriptions of how they train and prepare the pupils. As model examples they may get compositions or texts by professional writers to analyse and discuss. The teachers carry on such analysing exercises as a rule once or twice during the term to help the students learn for example how to structure a comparative or argumentative account. Especially the upper secondary school teachers use different exercises dealing with how to treat a certain type of text and how to make use of reference literature.

Much more often the teachers let the pupils work on exercises in writing technique and linguistic variation. Linguistic skill training is used more in the compulsory school than in the upper secondary school and then especially at the intermediate stage.

The training of the pupils' writing skill by means of exercises in arrangement of material and with the help of linguistic skill exercises is naturally a preparation for composition writing. This preparatioin is however more general and as a rule not aimed at the next writing occasion.

Swedish teachers at the senior level and the upper secondary school hardly at all prepare essay writing in he sense that they discuss with the students what to write the next time. They do not talk about how to plan the topic and they do not let the pupils have examples of words and expressions to be used. It is usually considered to be intermediate level methodology to prepare the pupils in this way, and 1/3 of the teachers on this level also say that they generally use such methods.

Also topics based on the students' own experiences or imagination may however need to be prepared. The content of such compositions can be presented in different ways, and especially at higher levels the students ought to get instruction in not only using chronological narration. For the younger pupils it is natural to relate chronologically, but the older students should also get an opportunity of learning to sort and group their thoughts and experiences in different ways. At all levels it is well-known that the pupils prefer to write about topics that are based on free imagination or their own experiences (Löfqvist, 1988b).

To make surveys, comparisons or analyses are examples of written tasks that are considered more difficult and for which the students ought to get assistance. Such topics can be treated in different ways depending on the type of material. But it should be important to give the students examples of how to tackle such topics. It is of use to them to know how they could weigh the pros and cons in an argumentative essay, even if a given model is not always the most suitable to use.

The more attention that is paid to preparations for the individual written task the minor the risks are that the pupils make poor performances. It is no doubt better to let the pupils get help and in that way succeed, than that they should be told afterwards how they ought to have solved the problem.

To prepare the individual compositions takes time, however, and many teachers probably think that the number of lessons in Sweden does not give space for much more than revision work. Especially at the senior level and in the upper secondary school the discussions seem to have been limited to follow-up. It may however be questioned if it is rational to use such long writing times as are usually offered. The longer writing times are most often justified by the assertion that the students need time to choose a topic, think it over and make preparations before they begin writing. But if the preparations have been made earlier it would be possible to limit writing time a great deal.

It seems to be common that teachers in other countries prepare the students for essay writing by discussing the topics with them. This is what happens in countries where the students could be either younger or somewhat older than Swedish senior level pupils in grade 9. As a natural consequence of this it is therefore also unusual that the students write about unprepared topics - the method which is the most common one in Sweden.

In Swedish schools the students usually get a number of topics or headlines to choose from, when they are to write essays. That is the most common basis that the students have to start out from. The teachers have not discussed them with the students or in any other way prepared the topics. Most often the students are not told about them beforehand and cannot therefore prepare themselves, e.g. by collecting information on the topics. On separate occasions the students get texts as a stimulus, and pictures could above all in the compulsory school excite imagination.

Standardized achievement tests in the upper secondary school and similar test in the compulsory school are based on text material that the students may use when they write their own compositions. This way of stimulating writing is applied in many places, and on those occasions tests that have been given earlier are used for training. Trying in different ways to stimulate students and enrich them with views on the topic and the subject field should in all likelihood be a better starting-point for a successful result than today's practice with unprepared writing situations.

It is common usage that the students while they are writing work on their own. Only at the intermediate level some teachers take an active part in the pupils' work and lend them a helping hand. At the senior stage and in the upper secondary school comments and corrections come mainly afterwards. Even such a way of working may be questioned. Is it not better, in the writing situation in question, to make it possible for the students to ask, to discuss design and to revise than only afterwards give opinions and corrections? In test situations, e.g. when they do standardized achievement tests, the students should naurally work without being helped while writing.

In the year 1988 the compulsory school got a new syllabus in Swedish about 3-4 years after the IEA study was made. It is, however, interesting in this connection, as it constitutes yet another step in the development of writing instruction as i can be understood in the syllabuses of 1962, 1969, 1980 and up to 1988 (Löfqvist, 1989a). Only a few comments will be presented.

With the 1980 curriculum for the compulsory school a different view of writing instruction was created. What was briefly prescribed in the main components of the curricula was given a more detailed description in supplementary comments, named Writing. In the 1980 syllabus these comments have been promoted to directions. Writing can be a process in many steps, in which the individual student with the support and suggestions from teachers and companions reaches the version that is the definite one and the one that is evaluated. In addition to this process described point by point functional writing with distinct purposes and readers is emphasized.

The curriculum gives a number of different examples of writing tasks that can be used to teach the students different types of writing. Most of these are also used in teaching. However, the students are rather seldom given an opportunity of practising certain of them, and some seem not to be trained at all during school hours either in the compulsory school or the upper secondary school.

The fact that certain types of writing are practised less or not at all must not necessarily mean deficiencies of education. If the students get a good general proficiency in writing, they are also most often able to cope with less usual writing situations. The correlations between their marks and the frequency of different tasks turned out to be low according to what was shown in the section called "Attitudes towards writing tasks". Irrespective of whether the students have practised a certain type of text or not, they are not at a loss unable to express themselves.

In the IEA project the students have been given writing tasks of various kinds. In connection with each item they have been asked to state if in their teaching the task is usual or unusual. Even as regards tasks where most of the students have considered it unusual they have managed well. In certain cases they have found it difficult to reach suitable style or linguistic tone, but they have not failed to deliver their message.

Even if it seems to be a fact that the students are also able to do unusual tasks, if they have general writing capacity, it is probably important to let them practise writing such tasks. There is undoubtedly greater variation in teaching if the students are placed in different writing situations. They are also to a larger extent given occasion to consider who they are writing for and how they should express themselves. Westman (1981) means that it seems as if school in a one-sided way emphasizes the explanatory, generalising essay on a neutral and not too controversial topic. That type of text is pracised by tradition but is not very well suited for "young people from 10 to 17 years of age as what it presupposes is maturity, experience and broad outlook, but hardly intelligence, imagination, frolic and freshness" (p 110). Even if explanatory and argumentative essays ought to remain several other kinds of text are better suited for children and young people, when they wish to express feelings and views in writing.

The older students usually write compositions a couple of times per term, whereas at the intermediate level composition writing most often occurs every second week. At the senior level and above all in the upper secondary school the writing situations take on the nature of tests. In the prescriptions of the Board of Education tests in written Swedish are mentioned and numbers and times for such tests are given. Writing times are generous, there are often for each term fixed writing dates and the compositions are usually scored. These writing sessions are looked upon as important, and it is natural for the students carefully to read through their essays and to correct them so that they can hand over a written product which is as satisfactory as possible.

The writing situation at the intermediate level is different. The pupils write more often and they get shorter writing time, often only one lesson. To write a composition is not looked upon as a test but rather as an exercise in telling about something that one has experienced or made up. The pupils say that they find it amusing to write but not that it is very important to read through and check what they have written.

What the pupils have written, irrespective of level, is most often only treated by the teacher in such a way that he reads the composition and, using correction marks, indicates what the pupils should change. In the compulsory school the teacher can discuss individually with the pupil what ought to be changed and improved. But no further revision or rewriting of the composition usually takes place. It is the praxis to consider something one has written as finished and completed, after principally formal mistakes that the teacher has pointed out have been corrected.

The pupils like to write. More than half of them find it amusing or very amusing, and only one out of ten think that it is boring. This interest should of course be taken care of by school.

It is to be wished that the writing sessions are more of exercises than tests. Written tests should naturally be arranged, when the students are placed in a situation when individully they have to write and finish a task, e.g. in the way that the standardized tests in the compulsory and the upper secondary school have been designed. But most of the writing sessions ought to be toned down. In the compulsory school curriculum a methodology is recommended which implies that the compositions should be revised by the students and teachers together. How work is to be modelled more in detail has been described in the supplementary comments Writing, and, as has been said above, it has been included as prescriptions for the teaching of writing in the new syllabus for Swedish 1988.

What the students have accomplished at the writing session should thus not be looked upon as finished. With the assistance of their companions and with the support of the teacher they should be able to revise their texts and improve them before these are finally evaluated.

8.2 Reliability of scoring

The results of scoring work lead to two conclusions. The first one refers to subjectivity in evaluation, which it has not been possible to cope with in spite of different measures, the second to the scoring method. The differences between scorers are great and it may therefore be hazardous to make comparisons between individuals or schools. We must be aware that the risk is considerable that an individual pupil's written performance should be estimated differently due to who his teacher is.

This state of things is of course unsatisfactory especially to the individual students, as the scoring of their achievements is often of importance to their continued activities, e.g. at admission into higher education. But probably we will not reach further as regards evaluation. One might compare with the critics' reviews of books, films, plays etc. Even if this comparison halts in certain respects, the deficiencies of agreement between critics just like among teachers are fundamentally due to differences of experiences, attitudes and views. We cannot therefore get to any standardized evaluations of artistic performances and they would not, one presumes, be desirable. Most often it is an average opinion that will guide the final estimate.

When trying to find ways to a better agreement between teachers' evaluations, we most often aim at hitting upon methods, that influence teachers to examine a solution of a writing task in the same way. We are concentrated on the product of the students' writing activities, and in an analysis of those we want to reach common scoring principles. Other ways have however been indicated in a manual of writing for American college students published recently (Purves et al, 1984). The pupils are requested to ask their teachers what they consider to be imporant to concentrate on during writing sessions, what school gives priority to at evaluations etc. Such questions may be put during or outside lessons but it is essential that the students should try to find out what the teachers consider worth thinking of if a student wants his composition to be evaluated as really good. In their writing the students should try and adjust themselves to their teachers' expectations.

It is possible that such a way of acting might result in greater agreement of evaluation within a school. It may give stimulus to a discussion of a common practice at school implying that it ought to be pointed out to the students what they should think of and observe while they are writing. At present the students seem to have a peculiar opinion about what the teachers consider most important to take into account in their estimate. Both in grade 9 and in the 2-year upper secondary school the students believe that the formal aspects are decisive of the composition score. If a student could only spell correctly, use full stops and write neatly he would have fulfilled the most important requirements. They carry greater weight than aspect of contents such as choice of subject, planning and imagination. The first to emphasize aspects of contents are the students of the 3-year upper secondary school lines, but even they attach great importance to the formal design (Löfqvist, 1988b).

That the students are of the opinion that their formal linguistic competence mainly determines what scores their compositions should have is something that they cannot in all likelihood have got from their teachers. It is true that the use and construction of language is practised in the compulsory school, most often in different separate exercises, but great importance is also attached to discussions of different types of texts and their structure. Probably discussions about planning and organization of the contents of compositions are not neglected in teaching. Rather the students' opinion is based on the view that their breaches of grammatical and ortographical conventions are chiefly the things that are marked in their compositions. The signs in the margin are more conspicuous and evident proofs of deficiencies than more general comments at the end of the composition.

The students views can be compared with what Hultman and Westman (1977) have observed on the compositions that were sent to the Board of Education after the standardized national tests (pp 227-228):

> "The corrections are interesting because they give an illustration of the grading technique of Swedish schools. It is principally the mistakes that are marked. The good points are almost never accentuated in the margin, but sometimes in a comprehensive estimate after the composition.By contrast with the negative comments - which are directed towards concrete phenomena in the text - such estimates tend to be sweeping and general. The students might easily get the impression that it is only the mistakes that count."

The students' view are further reinforced by the fact that the writing sessions at the senior level and in the upper secondary school which take place relatively seldom are looked upon as tests, and what the student has accomplished then is considered as the final product. The method which is nowadays prescribed for Swedish in the compulsory school implying that the students' written products are allowed gradually to develop in a discussion partly between the students themselves, partly between students and teachers, will - it is to be hoped - emphasize to the students in a stronger way the aspects of contents. The students get opportunities during the writing process of correcting formal blurs and mistakes, but above all they are allowed to revise and supplement what they have to say. A writing methodology like that does of course not minimize the importance of the capability of expressing oneself in a correct language but emphasizes to the students the weight which during evaluation is attached not only to ability to express oneself formally correctly but also to their skill at structuring their account.

Subjectivity of evaluation also implies that intentions of international comparisons in this IEA study cannot be carried into effect. The expectations which during the initial stage of the project were linked with the thought that by means of benchmark set and calibration essays it would be possible to neutralize national deviations have not been fulfilled. To make international comparisons seems chiefly to be obstructed by the problems that are combined with the contruction of instruments that can define an international scoring scale. It may be, however, that by means of a more specified plan it would have been possible to overcome the difficulties.

The study within the project that involved an investigation of whether the length of the students' composition as regards the number of words gives a result which is equivalent to the teachers' estimate gave a positive outcome. It was made clear that with the same security it is possible to evaluate students' written performances by counting the number of words in their compositions as by using two or three independent scorers. A mechanical language measure neglects other aspects at the evaluation but shows that in studies of this kind costs and efforts can be considerably reduced.

The method of using number of words as a language measure functions provided that the students do not know that the estimate is carried out in this mechanical way. It is therefore not applicable in the daily routine work at school. One can also assume that it would not be decisive in cases where the students get writing tasks such as introductions, summaries with a fixed number of words, résumés or instructions. Björnsson (1980) however found that even in summary exercises in grade 4 there were considerable variations in length. The summaries that received the highest scores were twice as long as the ones with the lowest scores.

8.3 Scoring method

The usual way of evaluating compositions in Sweden is to make an overall estimate, which is expressed with a figure, a letter or a standardized expression. The method used in the project of combining an overall estimate with an analytical marking of partial aspects is not applied in daily work. This more complicated form of estimate had as its purpose partly to diminish subjectivity of evaluation, partly to indicate a profile for a writing task in the sense that certain partial aspects would appear as easier or more difficult to master for the students.

Profiles are hardly possible to perceive in the IEA writing tasks. Only between the aspects content and disposition on the one hand and the formal writing competence on the other can one find a weaker correlation than between the others. The correlations between overall impression and partial aspects are however so high that during a continued analysing work it is sufficient to use the overall impression score as a measure of the students' capabilities. The second essential conclusion is thus that an analytical scoring in studies of this kind does not seem to add any qualities that cannot equally well be found only by means of an overall estimate.

But an analytical marking can have a considerable pedagogical value. It offers a possibility for the teachers more strongly than otherwise to emphasize a certain aspect of the students' handling of a writing task. It might be suitable in addition to the traditional overall estimate to complete e.g. with a disposition score, in the cases when the very organization of material can be varied. On other occasions attention may have to be directed especially towards style and linguistic tone, as in some of the writing tasks which are included in this study. Not least important it probably is to grade with a separate mark the legibility and neatness of the students' work.

Probably the students pay more attention to the evaluation of an aspect which is not only generally verbal but which is also expressed in a graded estimate. Such a more concrete and easily comprehensible evaluation can probably also help to explain to the students that it is not formal linguistic correctness that is decisive to the quality of what has been written.

References

Allwood, J. *Linguistic communication as action and cooperation. A study in pragmatics.* Göteborg: The university of Göteborg. Dept. of Linguistics, 1976.

Arnman, G. & Jönsson, I. *Segregation och svensk skola.* /Segregation and Swedish school./ Lund: Studentlitteratur, 1983.

Baker, E.L. The specification of writing tasks. In: Purves, A.C. & Takala, S. (Eds.) *An international perspective on the evolution of written composition.* Evaluation in Education: An international review series, Vol. 5, No. 3. Oxford: Pergamon, 1982. Pp 291-297.

Bereiter, C. Development in writing. In: Gregg, L.W. & Steinberg, E.R. (Eds.) *Cognitive processes in writing.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1980.

Bereiter, C. & Scardamalia, M. From conversation to composition: The role of instruction in a developmental process. In: Glaser, R. (Ed.) *Advances in instruction psychology.* Vol. 2. Hillsdale, NJ: Erlbaum, 1981.

Bernstein, B. & Henderson, D. Social class differences in the relevance of language to socialization. *Sociology*, 1969, *3*. Pp 1-20.

Bernstein, B. & Lundgren, U. *Makt, kontroll och pedagogik: studier av den kulturella reproduktionen.* /Power, control and education: studies of cultural reproduction./ Stockholm: Liber, 1983.

Björnsson, C.H. *Uppsatsbedömning och uppsatsskrivning.* /Scoring and writing compositions./ Stockholm: Almqvist & Wiksell, 1960.

Boalt, G. *Skolutbildning och skolresultat för barn ur olika samhällsgrupper i Stockholm.* /School training and school results of children from different social classes in Stockholm./ Stockholm: P.A. Nordstedt & Söner, 1947.

Boalt, G. & Husén, T. *Skolans sociologi.* /The sociology of the school./ Stockholm: Almqvist & Wiksell, 1967.

Brewer, W.F. Literacy theory, rhetoric and stylistics: Implications for psychology. In: Spiro, R.J., Bruce, B.C. & Brewer, W.F. (Eds.) *Theoretical issues in reading comprehension.* Hillsdale,NJ: Lawrence Erlbaum Associates, 1980.

Bühler, K. *Sprachteori.* /Theory of language./ Jena: Gustav Fischer Verlag, 1934.

Coffman, W.E. Essay examinations. In: Thorndike, R.L. (Ed.) *Educational measurement.* Washington, D.C.: American Council on Education, 1971. Pp 71-302.

Collins, A. & Gentner, D. A framework for a cognitive theory of writing. In: Gregg, L. & Steinberg, E. (Eds.) *Process in writing.* Hilldale, NJ: Lawrence Erlbaum Associates, 1980.

Eneskär, B. Forskning kring standardproven i svenska 1983-1986. /Research about the standardized tests in Swedish 1983-1986./ *Pedagogisk-psykologiska problem* (Malmö: Lärarhögskolan), Nr 492, 1988.

Garme, B. *Text och tanke.* /Text and thought./ Malmö: Liber, 1988.

BEST COPY AVAILABLE

Garrett, H. *Statistics in psychology and education*. London: Longmans Green and Co, 1958.

de Glopper, K. The results of the international scoring sessions. In: Gorman, T.P., Purves, A.C. & Degenhart, R.E. (Eds.) *The IEA study of written composition I: The international writing tasks and scoring scales*. NY: Pergamon, 1988. Pp 59-75.

Gorman, T.P., Purves, A.C. & Degenhart, R.E. (Eds.) *The IEA study of written composition I: The international writing tasks and scoring scales*. NY: Pergamon, 1988.

Grice, H.P. Meaning. *Philosophical Review*, 1957, *66*, 377-388.

Grogarn, M. *Dålig läsning: bakgrund, läsförmåga och läsintresse hos elever på Fo- och Ve-linjerna i gymnasieskolan*. /Bad reading: background, ability and interest in reading of students in the motor engineering line and the workshop line of the upper secondary school./ Lund: Liber, 1979.

Grundin, H. *Läs- och skrivförmågans utveckling genom skolåren*. /The development of reading and writing ability through the school years./ SÖ, Utbildningsforskning, Nr 20, 1975.

Guilford, J.P. *Fundamental statistics in psychology and education*. NY: Mc Graw-Hill, Inc., 1956.

Hansson, G. *Svensk skola i internationell belysning II. Läsning och litteratur*. /The Swedish school in an international light II. Reading and literature./ Stockholm: Almqvist & Wiksell, 1975.

Hayes, J.R. & Flower, L.S. Identifying the organization of writing processes. In: Gregg, L.W. & Steinberg, E.R. (Eds.) *Cognitive processes in writing*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1980. Pp 3-30.

Hersvall, M., Lindell, E. & Pettersson, I.-L. Om kvalitet i gymnasisters skriftspråk: Rapport från projektet "Svenska mått". /On quality in the written language of upper secondary school students./ *Pedagogisk-psykologiska problem* (Malmö: Lärarhögskolan), Nr 253, 1974.

Hultman, T.G. & Westman, M. Gymnasistsvenska. /Senior high school Swedish./ Lund: Liber Läromedel, 1977.

Hunt, K. *Differences in grammatical structures written at three grade levels, the structures to be analyzed by transformational methods*. U.S. Department of Health, Education and Welfare, 1964.

Härnqvist, K. & Grahm, Å. *Vägen genom gymnasiet. Elevernas syn på valsituationer och studieformer*. /The way through the upper seconday school. The students' ideas about elective situations and types of studies./ SOU 1963:15.

Jacobson, R. Linguistics and poetics. In: Sebeck, T.A. (Ed.) *Style in language*. New York: Wiley, 1960.

Jönsson, A. *TV - ett hot eller en resurs för barn?* /TV - a threat or a complement to school?/ Studia psychologica et pedagogica, Series altera LXXVI. Malmö: Gleerups, 1985.

Lamb, H. & Purves, A.C. The reflective composition. In: Gorman, T.P., Purves, A.C. & Degenhart, R.E. (Eds.) *The IEA study of written composition I: The international writing tasks and scoring scales*. NY: Pergamon, 1988. Pp 172-188.

Larsson, K. *Skrivförmåga. Studier i svenskt elevspråk*. /Writing ability. Studies of the language of Swedish school children./ Lund: Liber, 1984.

Lindell, E. Six reports on free writing: A summary of the FRIS project. *Didakometry* (Malmö, Sweden: School of Education), No. 61, 1980.

Lindell, E., Lundqvist, B., Martinsson, A., Nordlund, B. & Pettersson, I.-L. *Om fri skrivning i skolan*. /On free writing at school./ Stockholm: Liber/Utbildningsförlaget, 1978.

Lohmöller, J.-B. *LVPLS 1.6 program manual: Latent variables path analysis with partial least squares estimation*. Forschungsbericht 81.04. Neubiberg: Hochschule der Bundeswehr München, 1981.

Lundquist, B. Språkskillnader mellan kön, hem och skolor. Studier inom projektet FRIS (Fri skrivning i skolan). /Sex, home and school differences in language ability. Studies within the FRIS project (Free writing at school)./ *Pedagogisk-psykologiska problem* (Malmö: Lärarhögskolan), Nr 348, 1978.

Löfgren, H. & Löfqvist, G. Relationer mellan elevers sociokulturella bakgrund, undervisningen i skolan och skrivförmåga. /The relationships between students' social background, instruction in school and students' writing achievement./ *Pedagogisk-psykologiska problem* (Malmö: Lärarhögskolan), Nr 513, 1989.

Löfqvist, G. *Studier av elevers skrivförmåga. Uppläggning och innehåll*. /Studies of the writing ability of the students. Design and realization./ Lund: Studentlitteratur, 1988. (a)

Löfqvist, G. *Elever skriver uppsats*. /Students write compositions./ Lund: Studentlitteratur, 1988. (b)

Löfqvist, G. *Lärare i svenska*. /Swedish teachers of the mother tongue./ Lund: Studentlitteratur, 1988. (c)

Löfqvist., G. Skrivundervisning i läroplan och praxis./The teaching of writing in curriculum and practice./ *Pedagogisk-psykologiska problem* (Malmö: Lärarhögskolan), Nr 506, 1989. (a)

Löfqvist, G. Elevers tankevärld. /The world of studens' ideas./ *Pedagogisk-psykologiska problem* (Malmö: Lärarhögskolan), Nr 507, 1989. (b)

Löfqvist, G. Lärare bedömer uppsatser. /Teachers score compositions./ *Pedagogisk-psykologiska problem* (Malmö: Lärarhögskolan), Nr 510, 1989. (c)

Löfqvist, G. Elevers bakgrund och skrivintresse. /Students' background and interest in writing./ *Pedagogisk-psykologiska problem* (Malmö: Lärarhögskolan), Nr 516, 1989. (d)

Löfqvist, G. Prestationer i fri skrivning - en orsaksanalys. /Achievements in free writing - a causal analysis./ *Pedagogisk-psykologiska problem* (Malmö: Lärarhögskolan), Nr 522, 1989. (e)

Löfqvist, G. IEA-undersökning av skrivförmåga: I. Så ser den svenska skolan ut - några viktiga kringfaktorer II. Svensk skola mer homogen än de flesta andra länders. /IEA study about writing ability./ *Särtryck och smâtryck* (Malmö: Lärarhögskolan), Nr 639, 1989. (f)

Martinsson, A. Fri skrivning i skolan (FRIS): Elever, stoff och språkanalys. /Free writing at school (FRIS): Pupils' essay themes and language analysis./ *Pedagogisk-psykologiska problem* (Malmö: Lärarhögskolan), Nr 348, 1978.

Page, E.B. & Paulus, D.H. *The analysis of essays by computer. Final report.* Washington, D.C.: Office of Education (DHEW), Bureau of Research, 1968.

Pettersson, I.-L. Stimulering inför fri skrivning i skolan. Attityder samt effekter av försök med varierande stimuleringsåtgärder. En studie inom projektet FRIS (Fri skrivning i skolan). /Stimulating free writing at school: Attidudes and effects of experiments with different stimuli. A study within the FRIS project (Free writing at school)./ *Pedagogisk-psykologiska problem* (Malmö: Lärarhögskolan), Nr 351, 1978. (a)

Pettersson, I.-L. Kreativ skrivning. En studie inom projektet FRIS (Fri skrivning i skolan). /Creative writing. A study within the FRIS project (Free writing at school)./ *Pedagogisk-psykologiska problem* (Malmö: Lärarhögskolan), Nr 352, 1978. (b)

Pettersson, Å. Nittonårssvenska. Språket i 1977 års centrala prov i svenska för gymnasiets årskurs 3. /Swedish at Nineteen. The language of written compositions, spring 1977./ *Pedagogisk-psykologiska problem* (Malmö: Lärarhögskolan), Nr 377, 1980.

Pettersson, Å. Hur gymnasister skriver. /How upper secondary school students write./ *Svensklärarföreningen*, Nr 184, 1982.

Pettersson, Å. Utvecklingslinjer och utvecklingskrafter i elevernas uppsatser. I: Sandqvist, C. & Teleman, U. (Red.) *Språkutveckling under skoltiden.* /The development of the language during the schooldays./ Lund: Studentlitteratur, 1989.

Purves, A.C., Gorman, T.P. & Takala, S. The development of the scoring scheme and scales.In: Gorman, T.P., Purves, A.C. & Degenhart, R.E. (Eds.) *The IEA study of written composition I: The international writing tasks and scoring scales.* NY: Pergamon, 1988. Pp 41-58.

Purves, A.C., Takala, S. & Crismore, A. *How to write well in college.* NY: HBJ, 1984.

Takala, S. Student views on writing in nine countries. In: Degenhart, R.E. (Ed.) *Assessment of student writing in an international context.* Jyväskylä: University of Jyväskylä. Publication series B. Theory into practice 9. 1987. Pp 65-94.

Takala, S. Origins of the international study of writing. In: Gorman, T.P, Purves, A.C. & Degenhart, R.E. (Eds.) *The IEA study of written composition I: The international writing tasks and scoring scales.* NY: Pergamon, 1988. Pp 3-14.

Takala, S. & Vähäpassi, A. *On the specification of the domain of writing.* Jyväskylä: Pedagogiska forskningsinstitutet, 1983.

Teleman, U. *Manual för grammatisk beskrivning av talad och skriven svenska.* /Manual for a grammatical description of spoken and written Swedish./ Lund: Studentlitteratur, 1974.

BEST COPY AVAILABLE

Thorén, I. Att bedöma uppsatser. /To score compositions./ *I: Modersmålslärarnas förening. Årsskrift 1951.* /The union of the teachers of the mother tongue in Sweden. Year-book 1951./

Thorndike, R.L. *Personnel selection.* NY: Wiley, 1959.

Westman, M. Om textnormer i skolan. I: Larsson, K. (Red.) *Elevsvenska.* /The students' language./ Lund: Studentlitteratur, 1981. Pp 104-117.

White, J. & Löfqvist, G. Pragmatic writing tasks. In: Gorman, T.P., Purves, A.C. & Degenhart, R.E. (Eds.) *The IEA study of written composition I: The international writing tasks and scoring scales.* NY: Pergamon, 1988. Pp 79-99.

Wold, H. Soft modelling: The basic design and some extensions. In: Jöreskog, K.G. & Wold, H. (Eds.) *Systems under indirect observation: Causality, structure, prediction, part* Z. Amsterdam: North-Holland, 1982. Pp 1-54.

Vähäpassi, A. The domain of school writing and the development of writing tasks. In: Gorman, T.P., Purves, A.C. & Degenhart, R.E. (Eds.) *The IEA study of written composition I: The international writing tasks and scoring scales.* NY: Pergamon, 1988. Pp 79-99.

Published by the University of Lund, the Malmö School of Education, and the Society of Psychology and Pedagogy:

STUDIA PSYCHOLOGICA ET PAEDAGOGICA - SERIES ALTERA

---

1. Edlund, Sven: Diskussionen om begåvningsurvalet under reformations- och stormaktstiden. I. 1947.
2. Agrell, Jan: Nyinrättade professurer inom filosofiska fakulteten i Lund under 1800-talets första hälft. 1949.
3. Smith, Gudmund: Psychological Studies in Twin Differences. 1949.
4. Husén, Torsten: Om innebörden av psykologiska mätningar. 1949.
5. Smith, Gudmund: Interpretations of Behaviour Sequences. 1952.
6. Johannesson, Ingvar: Studier av sociala relationer mellan barn i folkskoleklasser. 1954.
7. Kragh, Ulf: The Actual-Genetic Model of Perception-Personality. 1955.
8. Bjerstedt, Åke: Interpretations of Sociometric Choice Status. 1956.
9. Werdelin, Ingvar: The Mathematical Ability. Experimental and Factorial Studies. 1958.
10. Werdelin, Ingvar: Geometrical Ability and the Space Factors in Boys and Girls. 1961.
11. Borg, Gunnar A.V.: Physical Performance and Perceived Exertion. 1962.
12. Lindell, Ebbe: Den svenska rättskrivningsmetodiken. 1964.
13. Bokander, Ingvar: Stereoscopically Induced Perceptual Conflict between Facial Photographs. 1968.
14. Bjerstedt, Åke: Twelve Years of Educational and Psychological Research in Sweden. 1968.
15. Larsson, Bernt: Bayes Strategies and Human Information Seeking. 1968.
16. Bjerstedt, Åke: Educational Technology. Instructional Programming and Didakometry. 1972.
17. Löfgren, Horst: The Measurement of Language Proficiency. 1972.
18. Bierschenk, Bernhard: Självkonfrontation via intern television i lärarutbildningen. 1972.
19. Gestrelius, Kurt: Job Analysis and Determination of Training Needs. 1972.
20. Sjöbäck, Hans: The Psychoanalytic Theory of Defensive Processes. 1973.
21. Larsson, Inger: Individualized Mathematics Teaching. 1973.

22 Fredriksson, Lennart: The Use of Self-observation and Questionnaires in Job Analysis for the Planning Training. 1974.

23. Idman, Pekka: Equality and Democracy: Studies of Teacher Training. 1974.
24. Bjerstedt, Åke: Social Development and Training in School. 1974.
25. Alwall, Ellen: The Religious Trend in Secular English--Irish Schoolbooks 1850-1861 and 1871-1880. 1974.

27. Lindsten, Carin: Hembygdskunskap i årskurs 3. 1975.
28. Bjerstedt, Åke: Pedagogisk forskning i Sverige. En sexårsbibliografi. 1975.
29. Nordén, Kerstin: Psychological Studies of Deaf Adolescents. 1975.
30. Westerlundh, Bert: Aggression, Anxiety and Defence. 1976.
31. Jernryd, Elisabeth: "Optimal auktoritets- och propagandaresistens". 1976.
32. Bierschenk, Bernhard & Bierschenk, Inger: A System for a Computer-Based Content Analysis of Interview Data. 1976.
33. Wiechel, Lennart: Roller och rollspel. 1976.
34. Hedlund, Carl: Commissioned Army Officers. 1976.
35. Bergsten-Brucefors, Agneta: Intelligence Patterns and Their Relation to Social Background. 1976.
36. Bjerstedt, Åke: Explorations in Creativity. 1976.
37. Wetterström, Magnhild: Medinflytande i skolan. 1977.
38. Bierschenk, Bernhard: Stimulating Strategies of Interactive Behaviour. 1978.
39. Drakenberg, Margareth: Förändringar i prestationsprocessen. 1978.
40. Flinck, Rune: Correspondence Education Combined with Systematic Telephone Tutoring. 1978.
41. Nordberg, Astrid: Uppfostran under underkastelse. 1978.
42. Eneskär, Barbro: Children's Language at Four and Six. 1978.
43. Eriksson, Mona: Utbildningspraktik. Om utbildning med handledda klient- och patientkontakter på hög nivå. 1978.
44. Hagberg, Bo: Cognitive Impairment and Personality Change in Relation to Brain Dysfunction. 1978.
45. Leonardsson, Sigurd: Den franska grammatikens historia i Sverige. I. 1978.
46. Leonardsson, Sigurd: Histoire de la grammaire française en Suède. II. 1978.
47. Carlsson, Bertil: Två utbildningsprojekt. En redovisning av arbetet att profilera verksamheten vid en nystartad svensk högskola. 1979.
48. Lindholm, Lena-Pia: Pupils' Attitudes to Equality between the Sexes. 1980.
49. Maximilian, V. Alexander: Functional Changes in the Cortex during Mental Activation. 1980.

50. Carlström, Inge: Law and Justice Education. 1980.
51. Prohovnik, Isak: Mapping Brainwork. 1980.
52. Bååth, John A. Postal Two-Way Communication in Correspondence Education. 1980.
53. Fritzell, Christer: Teaching, Sience and Ideology. 1981.
54. Bjerstedt, Åke (Ed.): Swedish Educational Research in the Early Eighties. 1981.
55. Maini, S.M.: Explorations in the Psychology of the Research Process. 1981.
56. Wiechel, Anita: Olika personalgruppers åsikter om barn i förskola och på lågstadium. 1981.
57. Bjerstedt, Åke: Pedagogisk forskning i Sverige. En sexårsbibliografi. 1981.
58. Bratt, Ingar: Läromedel i engelska utgivna i Sverige 1850-1905. 1982.
59. Gran, Birgitta: Från förskola till grundskola. Villkor för barns utveckling i åldrarna kring skolstarten. 1982.
60. Sages, Roger: Perceptual Determinants of Choice Behaviour. 1982.
61. Torper, Ulf: Tidsramar, tidsanvändning och kunskapsutveckling i den svenska grundskolan. 1982.
62. Adler, Erella: Mother-Image Perception in Normal and Disturbed Boys. 1982.
63. Bjerstedt, Åke: Fourteen Years of Educational and Psychological Research in Sweden. 1982.
64. Nilsson, Alf: The Mechanisms of Defence Within a Developmental Frame of Reference. 1983.
65. Annerblom, Marie-Louise: Att förändra könsroller: Ett feministiskt perspektiv på diskussionen om könsroller på dagis. 1983.
66. Holmberg, Lena: Om en speciallärares vardag. Analys av en dagbok. 1983.
67. Skov, Poul: Værdinormer om skolen: Analyse af konfliktmuligheder og mulige udviklingslinjer i skolen ud fra empiriske undersøgelser. 1983.
68. Viborg, Gardar G.: Studies in the Development of Social-Behaviour Understanding. 1984.
69. Carlsson, Gunborg: Teater för barn: Tre åldersgruppers upplevelser av professionell teater. 1984.
70. Welwert, Claes: Läsa eller lyssna? 1984.
71. Klasson, Maj: Högskolebibliotek i förändring. 1984.
72. Sjöberg, Marie: På bräcklig is: Studium av livsfilosofin hos några konstnärer, konstbeskådare och konstengagerade. 1984.
73. Andersson, Gunvor: Små barn på barnhem. 1984.
74. Bjerstedt, Åke: Pedagogisk forskning i Sverige: En treårsbibliografi. 1984.
75. Eppler, Mats & Nelander, Bo: Kartläggning och omformning av den psykosociala arbetsmiljön på ett sjukhus: En arbetslivspedagogisk studie. 1984.
76. Jönsson, Annelis: TV – ett hot eller en resurs för barn? En longitudinell studie av relationen mellan TV och skola. 1985.
77. Berglund, Lars: Decentraliserat högstadium. En longitudinell studie av elevers skoltrivsel och prestationer samt föräldrars och lärares attityder i en decentraliserad högstadieorganisation. 1985.
78. Teasdale, Thomas W.: Familial Influences in Social Class, Educational Level and Intelligence. 1985.
79. Hellström, Leif: Undervisningsmetodisk förändring i matematik: Villkor och möjligheter. 1985.
80. Ritchey, Elisabeth: Pedagogisk teori och vetenskapspolitik. Tre historiska strömningar. 1985.
81. Elliott, Carla J.: Application of the Repertory Grid Model to the Study of the Psychoanalytic Concept of Sublimation. 1985.
82. Klingberg, Göte: Children's Fiction in the Hands of the Translators. 1986.
83. Briem, Valdimir: Behaviour Observation: Theory and Application in Psychological Research. 1986.
84. Bjurwill, Christer: Framtidsföreställningar: Analys och tolkning av 900 elevers uppsatser och semantiska skattningar. 1986.
85. Åkerberg, Hans: Livet som utmaning: Existentiell ångest hos svenska gymnasieelever. 1987.
86. Bjerstedts, Åke: Pedagogisk forskning i Sverige: En treårsbibliografi till. 1987.
87. Berglund, Brigitte: Pedagogiska dagböcker i lärarutbildningen. 1988.
88. Svensson, Bengt E.: Mot samlad skoldag. 1988.
89. Rosenqvist, Jerry: Särskolan i ett arbetsmarknadsperspektiv. Om lärares och handledares uppfattningar av arbete och elevernas framtid i arbetslivet. 1988.
90. Rivano-Fischer, Marcelo: Nonverbal Behaviour in Interpersonal Situations; Notes on Spatial Behaviour in Public Places. 1988.
91. Varming, O.: Holdninger til børn. 1988.
92. Ozolins, Andrejs R.: Defence Patterns and Non-Communicative Body Movements: An Experimental Study with the Defence Mechanism Test. 1989.
93. Löfqvist, Gert: The IEA Study of Written Composition in Sweden. 1990.